Invariance Principle on the Slice

Yuval Filmus[1], Guy Kindler[2], Elchanan Mossel[3], and Karl Wimmer[4]

1 **Technion – Israel Institute of Technology, Haifa, Israel**
yuvalfi@cs.technion.ac.il

2 **The Hebrew University of Jerusalem, Jerusalem, Israel**
gkindler@cs.huji.ac.il

3 **The Wharton School, University of Pennsylvania, Philadelphia, USA; and University of California, Berkeley, USA**
mossel@wharton.upenn.edu

4 **Duquesne University, Pittsburgh, USA**
wimmerk@duq.edu

Abstract

The non-linear invariance principle of Mossel, O'Donnell and Oleszkiewicz establishes that if $f(x_1, \ldots, x_n)$ is a multilinear low-degree polynomial with low influences then the distribution of $f(\mathcal{B}_1, \ldots, \mathcal{B}_n)$ is close (in various senses) to the distribution of $f(\mathcal{G}_1, \ldots, \mathcal{G}_n)$, where $\mathcal{B}_i \in_R \{-1, 1\}$ are independent Bernoulli random variables and $\mathcal{G}_i \sim \mathrm{N}(0, 1)$ are independent standard Gaussians. The invariance principle has seen many application in theoretical computer science, including the *Majority is Stablest* conjecture, which shows that the Goemans–Williamson algorithm for MAX-CUT is optimal under the Unique Games Conjecture.

More generally, MOO's invariance principle works for any two vectors of hypercontractive random variables $(\mathcal{X}_1, \ldots, \mathcal{X}_n), (\mathcal{Y}_1, \ldots, \mathcal{Y}_n)$ such that (i) *Matching moments*: $\mathcal{X}_i$ and $\mathcal{Y}_i$ have matching first and second moments, (ii) *Independence*: the variables $\mathcal{X}_1, \ldots, \mathcal{X}_n$ are independent, as are $\mathcal{Y}_1, \ldots, \mathcal{Y}_n$.

The independence condition is crucial to the proof of the theorem, yet in some cases we would like to use distributions $(\mathcal{X}_1, \ldots, \mathcal{X}_n)$ in which the individual coordinates are not independent. A common example is the uniform distribution on the *slice* $\binom{[n]}{k}$ which consists of all vectors $(x_1, \ldots, x_n) \in \{0, 1\}^n$ with Hamming weight k. The slice shows up in theoretical computer science (hardness amplification, direct sum testing), extremal combinatorics (Erdős–Ko–Rado theorems) and coding theory (in the guise of the Johnson association scheme).

Our main result is an invariance principle in which $(\mathcal{X}_1, \ldots, \mathcal{X}_n)$ is the uniform distribution on a slice $\binom{[n]}{pn}$ and $(\mathcal{Y}_1, \ldots, \mathcal{Y}_n)$ consists either of n independent $\mathrm{Ber}(p)$ random variables, or of n independent $\mathrm{N}(p, p(1-p))$ random variables. As applications, we prove a version of *Majority is Stablest* for functions on the slice, a version of Bourgain's tail theorem, a version of the Kindler–Safra structural theorem, and a stability version of the t-intersecting Erdős–Ko–Rado theorem, combining techniques of Wilson and Friedgut.

Our proof relies on a combination of ideas from analysis and probability, algebra and combinatorics. In particular, we make essential use of recent work of the first author which describes an explicit Fourier basis for the slice.

1998 ACM Subject Classification F.0 [Theory of Computation] General

Keywords and phrases analysis of boolean functions, invariance principle, Johnson association scheme, the slice

Digital Object Identifier 10.4230/LIPIcs.CCC.2016.15

31st Conference on Computational Complexity (CCC 2016).
Editor: Ran Raz; Article No. 15; pp. 15:1–15:10
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany
COMPUTATIONAL COMPLEXITY CONFERENCE

1 Introduction

Analysis of Boolean functions is an area at the intersection of theoretical computer science, functional analysis and probability theory, which traditionally studies Boolean functions on the Boolean cube $\{0,1\}^n$. A recent development in the area is the non-linear *invariance principle* of Mossel, O'Donnell and Oleszkiewicz [11], a vast generalization of the fundamental Berry–Esseen theorem. The Berry–Esseen theorem is a quantitative version of the Central Limit Theorem, giving bounds on the speed of convergence of a sum $\sum_i X_i$ to the corresponding Gaussian distribution. Convergence occurs as long as none of the summands X_i is too "prominent". The invariance principle is an analog of the Berry–Esseen theorem for low-degree polynomials. Given a low-degree polynomial f on n variables in which none of the variables is too prominent (technically, f has low *influences*), the invariance principle states that the distribution of $f(X_1,\dots,X_n)$ and $f(Y_1,\dots,Y_n)$ is similar as long as each of the vectors $(X_1,\dots,X_n)$ and $(Y_1,\dots,Y_n)$ consists of independent coordinates, the distributions of X_i, Y_i have matching first and second moments, and the variables X_i, Y_i are hypercontractive.

The invariance principle came up in the context of proving a conjecture, *Majority is Stablest*, claiming that the majority function is the most noise stable among functions which have low influences. It is often applied in the following setting: the X_i are skewed Bernoulli variables, and the Y_i are the matching normal distributions. The invariance principle allows us to analyze a function on the Boolean cube (corresponding to the X_i) by analyzing its counterpart in Gaussian space (corresponding to the Y_i), in which setting it can be analyzed using geometric methods. This approach has been used to prove many results in analysis of Boolean functions (see for example [8]).

The proof of the invariance principle relies on the product structure of the underlying probability spaces. The challenge of proving an invariance principle for non-product spaces seems far from trivial. Here we prove such an invariance principle for the distribution over $X_1,\dots,X_n$ which is uniform over the *slice* $\binom{[n]}{k}$, defined as:

$$\binom{[n]}{k} = \{(x_1,\dots,x_n) \in \{0,1\}^n : x_1 + \cdots + x_n = k\}.$$

This setting arises naturally in hardness of approximation, see e.g. [3], and in extremal combinatorics (the Erdős–Ko–Rado theorem and its many extensions).

Our invariance principle states that if f is a low-degree function on $\binom{[n]}{k}$ having low influences, then the distributions of $f(X_1,\dots,X_n)$ and $f(Y_1,\dots,Y_n)$ are close, where $X_1,\dots,X_n$ is the uniform distribution on $\binom{[n]}{k}$, and $Y_1,\dots,Y_n$ are either independent Bernoulli variables with expectation k/n, or independent Gaussians with the same mean and variance.

The classical invariance principle is stated only for low-influence functions. Indeed, high-influence functions like $f(x_1,\dots,x_n) = x_1$ behave very differently on the Boolean cube and on Gaussian space. For the same reason, the condition of low-influence is necessary when comparing functions on the slice and on Gaussian space.

The invariance principle allows us to generalize two fundamental results to this setting: Majority is Stablest and Bourgain's tail bound. Using Bourgain's tail bound, we prove an analog of the Kindler–Safra theorem, which states that if a Boolean function is close to a function of constant degree, then it is close to a junta.

As a corollary of our Kindler–Safra theorem, we prove a stability version of the t-intersecting Erdős–Ko–Rado theorem, combining the method of Friedgut [7] with calculations of Wilson [12]. Friedgut showed that a t-intersecting family in $\binom{[n]}{k}$ of almost maximal size $(1-\epsilon)\binom{n-t}{k-t}$ is close to an optimal family (a t-star) as long as $\lambda < k/n < 1/(t+1) - \zeta$

(when $k/n > 1/(t+1)$, t-stars are no longer optimal). We extend his result to the regime $k/n \approx 1/(t+1)$.

The classical invariance principle is stated for *multilinear* polynomials, implicitly relying on the fact that every function on $\{0,1\}^n$ can be represented (uniquely) as a multilinear polynomial, and that multilinear polynomials have the same mean and variance under any product distribution in which the individual factors have the same mean and variance. In particular, the classical invariance principle shows that the correct way to lift a low-degree, low-influence function from $\{0,1\}^n$ to Gaussian space is via its multilinear representation.

The analogue of the collection of low degree multilinear functions on the discrete cube is given by the collection of low degree multilinear polynomials annihilated by the operator $\sum_{i=1}^n \frac{\partial}{\partial x_i}$. Dunkl [4, 5] showed that every function on the slice has a unique representation as a multilinear polynomial annihilated by the operator $\sum_{i=1}^n \frac{\partial}{\partial x_i}$. We call a polynomial satisfying this condition a *harmonic function*. In a recent paper [6], the first author showed that low-degree harmonic functions have *similar* mean and variance under both the uniform distribution on the slice and the corresponding Bernoulli and Gaussian product distributions. This is a necessary ingredient in our invariance principle.

Our results also apply for function on the slice that are not written in their harmonic representation. Starting with an arbitrary multilinear polynomial f, there is a unique harmonic function $\tilde{f}$ agreeing with f on a given slice. We show that as long as f depends on few coordinates, the two functions f and $\tilde{f}$ are close as functions over the Boolean cube. This implies that f behaves similarly on the slice, on the Boolean cube, and on Gaussian space.

Our proof combines algebraic, geometric and analytic ideas. A coupling argument, which crucially relies on properties of harmonic functions, shows that the distribution of a low-degree, low-influence harmonic function f is approximately invariant when we move from the original slice to nearby slices. Taken together, these slices form a thin layer around the original slice, on which f has roughly the same distribution as on the original slice. The classical invariance principle implies that the distribution of f on the layer is close to its distribution on the Gaussian counterpart of the layer, which turns out to be *identical* to its distribution on all of Gaussian space, completing the proof.

A special case of our main result can be stated as follows.

▶ **Theorem 1.1.** *For every $\epsilon > 0$ and integer $d \geq 0$ there exists $\tau = \tau(\epsilon, d) > 0$ such that the following holds. Let $n \geq 1/\tau$, and let f be a harmonic multilinear polynomial of degree d such that with respect to the uniform measure ν_{pn} on the slice $\binom{[n]}{pn}$, the variance of f is at most 1 and all influences of f are bounded by τ.*

The CDF distance between the distribution of f on the slice ν_{pn} and the distribution of f under the product measure μ_p with marginals $\mathrm{Ber}(p)$ is at most ϵ: for all $\sigma \in \mathbb{R}$,

$$|\Pr_{\nu_{pn}}[f < \sigma] - \Pr_{\mu_p}[f < \sigma]| < \epsilon.$$

Subsequent to this work, the first and third author came up with an alternative proof of Theorem 1.1 [10] which doesn't require the influences of f to be bounded. The proof is completely different, connecting the measures μ_p and ν_{pn} directly without recourse to Gaussian space. While the main result of [10] subsumes the main result of this paper, we believe that both approaches have merit. Furthermore, the applications of the invariance principle appearing here are not reproduced in [10].

Paper organization

An overview of our main results and methods appears in Section 2. Some open problems are described in Section 3. All proofs have been relegated to the full version of the paper, available online at `http://arxiv.org/abs/1504.01689`.

2 Overview

The goal of this section is to provide an overview of the results proved in this paper and the methods used to prove them. It is organized as follows. Some necessary basic definitions appear in Subsection 2.1. The invariance principle, its proof, and some standard consequences are described in Subsection 2.2. Some applications of the invariance principle appear in Subsection 2.3: versions of Majority is stablest, Bourgain's theorem, and the Kindler–Safra theorem for the slice. An application of the Kindler–Safra theorem to extremal combinatorics is described in Subsection 2.4. Finally, Subsection 2.5 presents results for non-harmonic multilinear polynomials.

2.1 Basic definitions

Measures

Our work involves three main probability measures, parametrized by an integer n and a probability $p \in (0,1)$:

- μ_p is the product distribution supported on the Boolean cube $\{0,1\}^n$ given by $\mu_p(S) = p^{|S|}(1-p)^{n-|S|}$.
- ν_{pn} is the uniform distribution on the slice $\binom{[n]}{pn} = \{(x_1,\ldots,x_n) \in \{0,1\}^n : x_1+\cdots+x_n = pn\}$ (we assume pn is an integer).
- $\mathcal{G}_p$ is the Gaussian product distribution $\mathrm{N}((p,\ldots,p), p(1-p)I_n)$ on Gaussian space $\mathbb{R}^n$.

We denote by $\|f\|_\pi$ the L2 norm of the polynomial f with respect to the measure π.

Harmonic polynomials

As stated in the introduction, we cannot expect an invariance principle to hold for all multilinear polynomials, since for example the polynomial $x_1+\cdots+x_n-pn$ vanishes on the slice but not on the Boolean cube or on Gaussian space. We therefore restrict our attention to *harmonic* multilinear polynomials, which are multilinear polynomials f satisfying the differential equation

$$\sum_{i=1}^n \frac{\partial f}{\partial x_i} = 0.$$

(The name *harmonic*, whose common meaning is different, was lifted from the literature.)

Dunkl [4, 5] showed that every function on the slice $\binom{[n]}{pn}$ has a unique representation as a harmonic multilinear polynomial whose degree is at most $\min(pn,(1-p)n)$. This is the analog of the well-known fact that every function on the Boolean cube has a unique representation as a multilinear polynomial.

One crucial property of low-degree harmonic multilinear polynomials is invariance of their L2 norm: for any $p \le 1/2$ and any harmonic multilinear polynomial f of degree $d \le pn$,

$$\|f\|_{\mu_p} = \|f\|_{\mathcal{G}_p} = \|f\|_{\nu_{pn}}\left(1 \pm O\left(\frac{d^2}{p(1-p)n}\right)\right).$$

This is proved in Filmus [6], and in fact this result (and its applications in the present work) was the main motivation for [6].

Influences

The classical definition of influence for a function f on the Boolean cube goes as follows. Define $f^{[i]}(x) = f(x^{[i]})$, where $x^{[i]}$ results from flipping the ith coordinate of x. The ith *cube*-influence of f is given by

$$\mathrm{Inf}^c_i[f] = \|f - f^{[i]}\|^2_{\mu_p} = \left\|\frac{\partial f}{\partial x_i}\right\|^2_{\mu_p} = \frac{1}{p(1-p)} \sum_{i \in S} \hat{f}(S)^2.$$

This notion doesn't make sense for functions on the slice, since the slice is not closed under flipping of a single coordinate. Instead, we consider what happens when two coordinates are swapped. Define $f^{(ij)}(x) = f(x^{(ij)})$, where $x^{(ij)}$ results from swapping the ith and jth coordinates of x. The (i,j)th *slice*-influence of f is given by

$$\mathrm{Inf}^s_{ij}[f] = \mathop{\mathbb{E}}_{\nu_{pn}}[(f - f^{(ij)})^2].$$

The influence of a single coordinate i is then defined as

$$\mathrm{Inf}^s_i[f] = \frac{1}{n} \sum_{j=1}^n \mathrm{Inf}^s_{ij}[f].$$

The two definitions are related: in the complete version of the paper we show that if $d = O(\sqrt{n})$ then

$$\mathrm{Inf}^s_i[f] = O_p\left(\frac{d}{n}\mathbb{V}[f] + \mathrm{Inf}^s_c[f]\right).$$

(The variance can be taken with respect to either the Boolean cube or the slice, due to the L2 invariance property.)

Noise stability

The classical definition of noise stability for a function f on the Boolean cube goes as follows:

$$\mathbb{S}^c_\rho[f] = \mathbb{E}[f(x)f(y)],$$

where $x \sim \mu_p$ and y is obtained from x by letting $y_i = x_i$ with probability ρ, and $y_i \sim \mu_p$ otherwise.

The analogous definition on the slice is slightly more complicated. For a function f on the slice,

$$\mathbb{S}^s_\rho[f] = \mathbb{E}[f(x)f(y)],$$

where $x \sim \nu_{pn}$ and y is obtained from x by doing $\mathrm{Po}(\frac{n-1}{2}\log\frac{1}{\rho})$ random transpositions (here $\mathrm{Po}(\lambda)$ is a Poisson distribution with mean λ). That this definition is the correct analog can be seen through the spectral lens:

$$\mathbb{S}^c_\rho[f] = \sum_d \rho^d \|f^{=d}\|^2_{\mu_p}, \qquad \mathbb{S}^s_\rho[f] = \sum_d \rho^{d - d(d-1)/n} \|f^{=d}\|^2_{\mu_{pn}}.$$

Here $f^{=d}$ is the dth homogeneous part of f consisting of all monomials of degree d.

2.2 Invariance principle

Our main theorem is an invariance principle for the slice.

▶ **Theorem 2.1.** *Let f be a harmonic multilinear polynomial of degree d such that with respect to ν_{pn}, $\mathbb{V}[f] \le 1$ and $\mathrm{Inf}^s_i[f] \le \tau$ for all $i \in [n]$. Suppose that $\tau \le I_p^{-d}\delta^K$ and $n \ge I_p^d/\delta^K$, for some constants I_p, K. For any C-Lipschitz functional ψ and for $\pi \in \{\mathcal{G}_p, \mu_p\}$,*

$$|\mathop{\mathbb{E}}_{\nu_{pn}}[\psi(f)] - \mathop{\mathbb{E}}_{\pi}[\psi(f)]| = O_p(C\delta).$$

Proof sketch. Let ψ be a Lipschitz functional and f a harmonic multilinear polynomial of unit variance, low slice-influences, and low degree d. A simple argument (mentioned above) shows that f also has low cube-influences, and this implies that

$$\mathop{\mathbb{E}}_{\nu_k}[\psi(f)] \approx \mathop{\mathbb{E}}_{\nu_{pn}}[\psi(f)] \pm O_p\left(\frac{|k-np|}{\sqrt{n}} \cdot \sqrt{d}\right).$$

The idea is now to apply the multidimensional invariance principle jointly to f and to $S = \frac{x_1+\cdots+x_n-np}{\sqrt{p(1-p)n}}$, deducing

$$\mathop{\mathbb{E}}_{\mu_p}[\psi(f)\mathbf{1}_{|S|\le\sigma}] = \mathop{\mathbb{E}}_{\mathcal{G}_p}[\psi(f)\mathbf{1}_{|S|\le\sigma}] \pm \epsilon.$$

Let $\gamma_{p,q}$ be the restriction of $\mathcal{G}_p$ to the Gaussian slice $\{(x_1,\ldots,x_n) \in \mathbb{R}^n : x_1+\cdots+x_n = qn\}$. An easy argument shows that since f is harmonic, the distribution of $f(\mathcal{G}_p)$ and $f(\gamma_{p,q})$ is identical, and so

$$\mathop{\mathbb{E}}_{\mathcal{G}_p}[\psi(f)\mathbf{1}_{|S|\le\sigma}] = \Pr_{\mathcal{G}_p}[|S| \le \sigma] \mathop{\mathbb{E}}_{\mathcal{G}_p}[\psi(f)].$$

Similarly,

$$\mathop{\mathbb{E}}_{\mu_p}[\psi(f)\mathbf{1}_{|S|\le\sigma}] = \Pr_{\mu_p}[|S| \le \sigma](\mathop{\mathbb{E}}_{\mu_p}[\psi(f)] \pm O_p(\sigma\sqrt{d})).$$

Since $\Pr_{\mathcal{G}_p}[|S| \le \sigma] \approx \Pr_{\mu_p}[|S| \le \sigma] = \Theta_p(\sigma)$, we can conclude that

$$\mathop{\mathbb{E}}_{\nu_{pn}}[\psi(f)] \approx \mathop{\mathbb{E}}_{\mathcal{G}_p}[\psi(f)] \pm O_p\left(\sigma\sqrt{d} + \frac{\epsilon}{\sigma}\right).$$

By choosing σ appropriately, we balance the two errors and obtain our invariance principle. ◀

As corollaries, we bound the Lévy and CDF distances between $f(\nu_{pn})$, $f(\mu_p)$ and $f(\mathcal{G}_p)$:

▶ **Corollary 2.2.** *Let f be a harmonic multilinear polynomial of degree d such that with respect to ν_{pn}, $\mathbb{V}[f] \le 1$ and $\mathrm{Inf}^s_i[f] \le \tau$ for all $i \in [n]$. There are parameters X_p, X such that for any $0 < \epsilon < 1/2$, if $\tau \le X_p^{-d}\epsilon^X$ and $n \ge X_p^d/\epsilon^X$ then the Lévy distance between $f(\nu_{pn})$ and $f(\pi)$ is at most ϵ, for $\pi \in \{\mathcal{G}_p, \mu_p\}$. In other words, for all σ,*

$$\Pr_{\nu_{pn}}[f \le \sigma - \epsilon] - \epsilon \le \Pr_{\pi}[f \le \sigma] \le \Pr_{\nu_{pn}}[f \le \sigma + \epsilon] + \epsilon.$$

▶ **Corollary 2.3.** *Let f be a harmonic multilinear polynomial of degree d such that with respect to ν_{pn}, $\mathbb{V}[f] = 1$ and $\mathrm{Inf}^s_i[f] \le \tau$ for all $i \in [n]$. There are parameters Y_p, Y such that for any $0 < \epsilon < 1/2$, if $\tau \le (Y_p d)^{-d}\epsilon^{Yd}$ and $n \ge (Y_p d)^d/\epsilon^{Yd}$ then the CDF distance between $f(\nu_{pn})$ and $f(\pi)$ is at most ϵ, for $\pi \in \{\mathcal{G}_p, \mu_p\}$. In other words, for all σ,*

$$|\Pr_{\nu_{pn}}[f \le \sigma] - \Pr_{\pi}[f \le \sigma]| \le \epsilon.$$

The proofs of these corollaries closely follows the proof of the analogous results in [11].

2.3 Applications

As applications to our invariance principle, we prove analogues of three classical results in analysis of Boolean functions: Majority is stablest; Bourgain's theorem; and the Kindler–Safra theorem:

▶ **Theorem 2.4.** *Let $f\colon \binom{[n]}{pn} \to [0,1]$ have expectation μ and satisfy $\mathrm{Inf}^s_i[f] \le \tau$ for all $i \in [n]$. For any $0 < \rho < 1$, we have*

$$\mathbb{S}^s_\rho[f] \le \Gamma_\rho(\mu) + O_{p,\rho}\left(\frac{\log\log\frac{1}{\alpha}}{\log\frac{1}{\alpha}}\right) + O_\rho\left(\frac{1}{n}\right), \text{ where } \alpha = \min(\tau, \tfrac{1}{n}),$$

where $\Gamma_\rho(\mu)$ is the probability that two ρ-correlated Gaussians be at most $\Phi^{-1}(\mu)$ (here Φ is the CDF of a standard Gaussian).

▶ **Theorem 2.5.** *Fix $k \ge 2$. Let $f\colon \binom{[n]}{pn} \to \{\pm 1\}$ satisfy $\mathrm{Inf}^s_i[f^{\le k}] \le \tau$ for all $i \in [n]$. For some constants $W_{p,k}, C$, if $\tau \le W_{p,k}^{-1}\,\mathbb{V}[f]^C$ and $n \ge W_{p,k}/\,\mathbb{V}[f]^C$ then*

$$\|f^{>k}\|^2 = \Omega\left(\frac{\mathbb{V}[f]}{\sqrt{k}}\right).$$

▶ **Theorem 2.6.** *Fix the parameter $k \ge 2$. Let $f\colon \binom{[n]}{pn} \to \{\pm 1\}$ satisfy $\|f^{>k}\|^2 = \epsilon$. There exists a function $h\colon \binom{[n]}{pn} \to \{\pm 1\}$ of degree k depending on $O_{k,p}(1)$ coordinates (that is, invariant under permutations of all other coordinates) such that*

$$\|f - h\|^2 = O_{p,k}\left(\epsilon^{1/C} + \frac{1}{n^{1/C}}\right),$$

for some constant C.

The proof of Theorem 2.4 closely follows its proof in [11]. The proofs of the other two theorems closely follows analogous proofs in [9].

2.4 t-Intersecting families

As an application of our Kindler–Safra theorem, we prove a stability result for t-intersecting families.

First, a few definitions:

- A t-intersecting family $\mathcal{F} \subseteq \binom{[n]}{k}$ is one in which $|A \cap B| \ge t$ for any $A, B \in \mathcal{F}$.
- A t-star is a family of the form $\{A \in \binom{[n]}{k} : A \supseteq J\}$, where $|J| = t$.
- A $(t,1)$-Frankl family is a family of the form $\{A \in \binom{[n]}{k} : |A \cap J| \ge t+1\}$, where $|J| = t+2$.

Ahlswede and Khachatrian [1, 2] proved that if $n > (t+1)(k-t+1)$ and $\mathcal{F}$ is an intersecting family, then $|\mathcal{F}| \le \binom{n-t}{k-t}$, and furthermore equality holds if and only if $\mathcal{F}$ is a t-star. They also proved that when $n = (t+1)(k-t+1)$ the same upper bound holds, but now equality holds for both t-stars and $(t,1)$-Frankl families.

A corresponding stability result was proved by Friedgut [7]:

▶ **Proposition 2.7** (Friedgut)**.** *Let $t \ge 1$, $k \ge t$, $\lambda, \zeta > 0$, and $\lambda n < k < (\frac{1}{t+1} - \zeta)n$. Suppose $\mathcal{F} \subseteq \binom{[n]}{k}$ is a t-intersecting family of measure $|\mathcal{F}| = \binom{n-t}{k-t} - \epsilon\binom{n}{k}$. Then there exists a family $\mathcal{G}$ which is a t-star such that*

$$\frac{|\mathcal{F} \triangle \mathcal{G}|}{\binom{n}{k}} = O_{t,\lambda,\zeta}(\epsilon).$$

Friedgut's theorem requires k/n to be bounded away from $1/(t+1)$. Using the Kindler–Safra theorem on the slice rather than the Kindler–Safra theorem on the Boolean cube (which is what Friedgut uses), we can do away with this limitation:

▶ **Theorem 2.8.** *Let $t \geq 2$, $k \geq t+1$ and $n = (t+1)(k-t+1)+r$, where $r > 0$. Suppose that $k/n \geq \lambda$ for some $\lambda > 0$. Suppose $\mathcal{F} \subseteq \binom{[n]}{k}$ is a t-intersecting family of measure $|\mathcal{F}| = \binom{n-t}{k-t} - \epsilon\binom{n}{k}$. Then there exists a family $\mathcal{G}$ which is a t-star or a $(t,1)$-Frankl family such that*

$$\frac{|\mathcal{F} \triangle \mathcal{G}|}{\binom{n}{k}} = O_{t,\lambda}\left(\max\left(\left(\frac{k}{r}\right)^{1/C}, 1\right)\epsilon^{1/C} + \frac{1}{n^{1/C}}\right),$$

for some constant C.

Furthermore, there is a constant $A_{t,\lambda}$ such that $\epsilon \leq A_{t,\lambda}\min(r/k,1)^{C+1}$ implies that $\mathcal{G}$ is a t-star.

Our proof closely follows the argument of Friedgut [7], transplanting it from the setting of the Boolean cube to the setting of the slice, using calculations of Wilson [12] in the latter setting. The argument involves certain subtelties peculiar to the slice.

2.5 Non-harmonic functions

All results we have described so far apply only to harmonic multilinear polynomials. We mentioned that some of these results trivially don't hold for some non-harmonic multilinear polynomials: for example, $\sum_{i=1}^n x_i - np$ doesn't exhibit invariance. This counterexample, however, is a function depending on all coordinates. In contrast, we can show that some sort of invariance does apply for general multilinear polynomials that depend on a small number of coordinates:

▶ **Theorem 2.9.** *Let f be a multilinear polynomial depending on d variables, and let $\tilde{f}$ be the unique harmonic multilinear polynomial agreeing with f on $\binom{[n]}{pn}$, where $d \leq pn \leq n/2$. For $\pi \in \{\mu_p, \mathcal{G}_p\}$ we have*

$$\|f - \tilde{f}\|_\pi^2 = O\left(\frac{d^2 2^d}{p(1-p)n}\right)\|f\|_\pi^2.$$

Proof sketch. Direct calculation shows that if ω is a Fourier character than

$$\|\omega - \tilde{\omega}\|_{\mu_p}^2 = \|\omega - \tilde{\omega}\|_{\mathcal{G}_p}^2 = O\left(\frac{d^2}{p(1-p)n}\right),$$

where $\tilde{\omega}$ is defined analogously to $\tilde{f}$.

We can assume without loss of generality that f depends only on the variables in $[d] = \{1,\dots,d\}$. Since $\tilde{f} = \sum_{S \subseteq [d]} \hat{f}(S)\tilde{\omega}_S$,

$$\|f - \tilde{f}\|_\pi^2 \leq 2^d \sum_{S \subseteq [d]} \hat{f}(S)^2 O\left(\frac{d^2}{p(1-p)n}\right) = O\left(\frac{d^2 2^d}{p(1-p)n}\right)\|f\|_\pi^2,$$

using the Cauchy–Schwartz inequality. ◀

The idea of the proof is to prove a similar results for Fourier characters for individual Fourier characters, and then to invoke the Cauchy–Schwartz inequality.

As a consequence, if we have a multilinear polynomial f depending on a small number of variables, its harmonic projection $\tilde{f}$ (defined as in the theorem) has a similar expectation, L2 norm, variance and noise stability. This implies, for example, that our Majority is stablest theorem is tight: the harmonic projection of the majority of a small number of indices serves as the tight example.

3 Open problems

Our work gives rise to several open questions.

1. Prove (or refute) an invariance principle comparing ν_{pn} and $\gamma_{p,p}$ for arbitrary (non-harmonic) multilinear polynomials.
2. Prove a tight version of the Kindler–Safra theorem on the slice (Theorem 2.6).
3. The uniform distribution on the slice is an example of a negatively associated vector of random variables. Generalize the invariance principle to this setting.
4. The slice $\binom{[n]}{k}$ can be thought of as a 2-coloring of $[n]$ with a given histogram. Generalize the invariance principle to c-colorings with given histogram.
5. The slice $\binom{[n]}{k}$ has a q-analog: all k-dimensional subspaces of $\mathbb{F}_q^n$ for some prime power q. The analog of the Boolean cube consists of all subspaces of $\mathbb{F}_q^n$ weighted according to their dimension. Generalize the invariance principle to the q-analog, and determine the analog of Gaussian space.

Acknowledgements. This paper started its life when all authors were members of a semester-long program on "Real Analysis in Computer Science" at the Simons Institute for Theory of Computing at U.C. Berkeley. The authors would like to thank the institute for enabling this work.

Y.F. would like to mention that this material is based upon work supported by the National Science Foundation under agreement No. DMS-1128155. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors, and do not necessarily reflect the views of the National Science Foundation. The bulk of the work on this paper was done while at the Institute for Advanced Study, Princeton, NJ.

E.M. would like to acknowledge the support of the following grants: NSF grants DMS 1106999 and CCF 1320105, DOD ONR grant N00014-14-1-0823, and grant 328025 from the Simons Foundation.

K.W. would like to acknowledge the support of NSF grant CCF 1117079.

References

1 Rudolf Ahlswede and Levon H. Khachatrian. The complete intersection theorem for systems of finite sets. *Eur. J. Comb.*, 18(2):125–136, 1997.

2 Rudolf Ahlswede and Levon H. Khachatrian. A pushing-pulling method: New proofs of intersection theorems. *Combinatorica*, 19(1):1–15, 1999.

3 Roee David, Irit Dinur, Elazar Goldenberg, Guy Kindler, and Igor Shinkar. Direct sum testing. In *ITCS 2015*, 2015.

4 Charles F. Dunkl. A Krawtchouk polynomial addition theorem and wreath products of symmetric groups. *Indiana Univ. Math. J.*, 25:335–358, 1976.

5 Charles F. Dunkl. Orthogonal functions on some permutation groups. In *Relations between combinatorics and other parts of mathematics*, volume 34 of *Proc. Symp. Pure Math.*, pages 129–147, Providence, RI, 1979. Amer. Math. Soc.

6 Yuval Filmus. An orthogonal basis for functions over a slice of the boolean hypercube. *Elec. J. Comb.*, 23(1):P1.23, 2016.

7 Ehud Friedgut. On the measure of intersecting families, uniqueness and stability. *Combinatorica*, 28(5):503–528, 2008. doi:10.1007/s00493-008-2318-9.

8 Subhash Khot. Inapproximability of NP-complete problems, discrete fourier analysis, and geometry. In *Proceedings of the International Congress of Mathematicians*, Hyderabad, India, 2010.

9 Guy Kindler, Naomi Kirshner, and Ryan O'Donnell. Gaussian noise sensitivity and Fourier tails, 2014. Manuscript.

10 Elchanan Mossel and Yuval Filmus. Harmonicity and invariance on slices of the Boolean cube. In *31st Conf. Comp. Comp.*, 2016.

11 Elchanan Mossel, Ryan O'Donnell, and Krzysztof Oleszkiewicz. Noise stability of functions with low influences: Invariance and optimality. *Ann. Math.*, 171:295–341, 2010.

12 Richard M. Wilson. The exact bound in the Erdős-Ko-Rado theorem. *Combinatorica*, 4:247–257, 1984.