Penn State University
CSE 564 Complexity of Combinatorial Problems

Instructor: Pei Wu

Module
17

Invariance principle and MIS

This lecture explains the conceptual bridge between two worlds: Gaussian space (where geometry and isoperimetry give sharp extremal inequalities) and the Boolean cube $\{-1,1\}^n$ (where we study low-influence Boolean functions). The motivating application is the *Majority Is Stablest* (MIS) theorem of Mossel–O'Donnell–Oleszkiewicz, which plays a central role in hardness of approximation (e.g. in the analysis of Max Cut and Unique Games reductions).

The story has two main ingredients. First, Borell's Gaussian noise-stability theorem characterizes half-spaces as extremal objects for Gaussian noise stability. Second, the *invariance principle* is a far-reaching generalization of the central limit theorem showing that low-degree, low-influence multilinear polynomials have (approximately) the same distribution under Rademacher inputs and Gaussian inputs, when tested against smooth functions. Combining these yields MIS on the Boolean cube.

17.1 Noise stability and the Majority Is Stablest theorem

We begin with the basic notions on the Boolean cube.

Definition 17.1 (ρ-correlated inputs). Fix $\rho \in [-1,1]$. A pair $(X,Y) \in \{-1,1\}^n \times \{-1,1\}^n$ is *ρ-correlated* if the coordinates are independent and for each $i \in [n]$,

$$\mathbb{P}[Y_i = X_i] = \frac{1+\rho}{2}, \qquad \mathbb{P}[Y_i \neq X_i] = \frac{1-\rho}{2}.$$

Equivalently, $\mathbb{E}[X_i] = \mathbb{E}[Y_i] = 0$ and $\mathbb{E}[X_iY_i] = \rho$.

Definition 17.2 (Noise stability on the cube). For a function $f : \{-1,1\}^n \to \mathbb{R}$, the *noise stability* at correlation ρ is

$$\mathrm{Stab}_\rho(f) := \mathbb{E}\big[f(X)f(Y)\big],$$

where (X,Y) is a ρ-correlated pair as above.

Definition 17.3 (Influence). For $f : \{-1,1\}^n \to \mathbb{R}$, the *influence* of coordinate i is

$$\mathrm{Inf}_i(f) \;:=\; \mathbb{E}\big[\mathrm{Var}(f(X) \mid X_1,\dots,X_{i-1},X_{i+1},\dots,X_n)\big] = \frac{1}{4}\mathbb{E}\big[(f(X)-f(X^{\oplus i}))^2\big],$$

where $X^{\oplus i}$ denotes X with the ith bit flipped.

The canonical low-influence Boolean function is majority.

Definition 17.4 (Majority). Define $\mathrm{MAJ}_n : \{-1,1\}^n \to \{-1,1\}$ by

$$\mathrm{MAJ}_n(x) \;:=\; \mathrm{sign}\left(\sum_{i=1}^n x_i\right),$$

with ties broken arbitrarily (e.g. output $+1$ when the sum is 0).

Claim 17.5 (Noise stability of majority converges to a Gaussian quantity). *For every fixed $\rho \in (-1,1)$,*

$$\lim_{n\to\infty} \mathrm{Stab}_\rho(\mathrm{MAJ}_n) \;=\; 1 - \frac{2\arccos(\rho)}{\pi} \;=\; \frac{2}{\pi}\arcsin(\rho).$$

Proof sketch (via CLT). Let (X,Y) be ρ-correlated. Then

$$\mathrm{Stab}_\rho(\mathrm{MAJ}_n) = \mathbb{E}\Big[\mathrm{sign}\Big(\sum_{i=1}^n X_i\Big)\cdot \mathrm{sign}\Big(\sum_{i=1}^n Y_i\Big)\Big].$$

Rescale by $\sqrt{n}$ and apply the multivariate CLT:

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n (X_i,Y_i) \;\Longrightarrow\; (G,G'),$$

where (G,G') is a centered bivariate normal with $\mathbb{E}[G^2] = \mathbb{E}[(G')^2] = 1$ and $\mathbb{E}[GG'] = \rho$. Therefore the stability converges to $\mathbb{E}[\mathrm{sign}(G)\mathrm{sign}(G')]$. A standard geometric computation for correlated Gaussians shows $\mathbb{P}[\mathrm{sign}(G) \neq \mathrm{sign}(G')] = \arccos(\rho)/\pi$, hence

$$\mathbb{E}[\mathrm{sign}(G)\mathrm{sign}(G')] = 1 - 2\mathbb{P}[\mathrm{sign}(G)\neq\mathrm{sign}(G')] = 1 - \frac{2\arccos(\rho)}{\pi}.$$

∎

The MIS theorem says that, among *balanced* low-influence functions, the noise stability is essentially pinned to the Gaussian half-space value (up to ε).

Theorem 17.6 (Majority Is Stablest (Mossel–O'Donnell–Oleszkiewicz)). *Fix $\rho \in (-1,0)$ and $\varepsilon > 0$. There exists $\delta = \delta(\rho,\varepsilon) > 0$ such that for every $f : \{-1,1\}^n \to [-1,1]$ satisfying*

$$\mathbb{E}[f] = 0 \qquad \textit{and} \qquad \max_{i\in[n]} \mathrm{Inf}_i(f) \le \delta,$$

we have

$$\mathrm{Stab}_\rho(f) \;\ge\; 1 - \frac{2\arccos(\rho)}{\pi} - \varepsilon.$$

(For $\rho \in (0,1)$ the same conclusion holds with the inequality reversed.)

REMARK 17.7. The right-hand side is exactly the limit noise stability of MAJ_n (equivalently, of a Gaussian half-space), and the theorem asserts that *no* balanced function with all influences small can beat it by more than ε.

17.2 Gaussian noise stability and Borell's theorem

The Gaussian analogue replaces $\{-1,1\}^n$ by $\mathbb{R}^n$ with the standard Gaussian measure.

DEFINITION 17.8 (ρ-correlated Gaussians). Let $G \sim \mathcal{N}(0, I_n)$ and let $H := \rho G + \sqrt{1-\rho^2}\, G'$, where $G' \sim \mathcal{N}(0, I_n)$ is independent of G. Then (G, H) is a pair of *ρ-correlated Gaussians* in $\mathbb{R}^n$.

Borell's theorem identifies half-spaces as extremal for Gaussian noise stability.

THEOREM 17.9 (Borell's isoperimetric (noise stability) theorem (1980s)). *Fix $\rho \in (-1,1)$ and let $f : \mathbb{R}^n \to [-1,1]$ satisfy $\mathbb{E}[f(G)] = 0$. Then*

$$\mathrm{Stab}_\rho(f) \begin{cases} \le 1 - \dfrac{\arccos(\rho)}{\pi}, & \rho > 0, \\ \ge 1 - \dfrac{\arccos(\rho)}{\pi}, & \rho < 0. \end{cases}$$

Moreover, equality (for $\rho \neq 0$) is attained by half-spaces $A = \{x : \langle u, x\rangle \ge 0\}$.

For a half-space of Gaussian measure 1/2, the probability that G and H fall on the same side is exactly $1 - \arccos(\rho)/\pi$. Translating to ± 1 functions via $f(x) = \mathrm{sign}(\langle u, x\rangle)$ recovers the earlier constant $1 - 2\arccos(\rho)/\pi$.

17.3 The invariance principle

The invariance principle can be viewed as a "nonlinear CLT" for low-degree, low-influence multilinear polynomials.

DEFINITION 17.10 (Multilinear polynomial and its restriction). A multilinear polynomial $P : \mathbb{R}^n \to \mathbb{R}$ has the form

$$P(z) \;=\; \sum_{S \subseteq [n]} \widehat{P}(S) \prod_{i \in S} z_i,$$

and its degree is $\deg(P) := \max\{|S| : \widehat{P}(S) \neq 0\}$. We write $P(X)$ for $X \in \{-1,1\}^n$ and $P(G)$ for $G \sim \mathcal{N}(0, I_n)$.

DEFINITION 17.11 (Influence of a multilinear polynomial). For a multilinear P, define

$$\mathrm{Inf}_i(P) \;:=\; \sum_{S \ni i} \widehat{P}(S)^2.$$

This coincides (up to standard conventions) with $\mathbb{E}[(\partial_i P(X))^2]$ under $X \sim \{-1,1\}^n$.

Theorem 17.12 (Invariance principle (informal form)). *Let $P : \mathbb{R}^n \to \mathbb{R}$ be multilinear of degree D and assume*

$$\max_{i \in [n]} \operatorname{Inf}_i(P) \le \delta.$$

Then for every $\varphi : \mathbb{R} \to \mathbb{R}$ that is four times differentiable with $\left\|\varphi^{(4)}\right\|_\infty < \infty$,

$$\left| \mathbb{E}_{X \sim \{-1,1\}^n}[\varphi(P(X))] - \mathbb{E}_{G \sim \mathcal{N}(0, I_n)}[\varphi(P(G))] \right| \le O\Big(D \cdot C^D \cdot \delta \cdot \left\|\varphi^{(4)}\right\|_\infty\Big),$$

for an absolute constant C (in the lecture notes one may take $C = 9$).

Remark 17.13. When $P(z) = \frac{1}{\sqrt{n}} \sum_{i=1}^n z_i$ (degree 1), this recovers the classical CLT statement that $\sum X_i/\sqrt{n}$ and $\sum G_i/\sqrt{n}$ have similar distributions under smooth tests. The power of the invariance principle is that it applies to *arbitrary* low-degree polynomials, as long as no single coordinate has too much influence on the value.

17.4 Borell + invariance ⇒ Majority Is Stablest

At a high level, the derivation of MIS has the following shape.

1. Start from a balanced Boolean function $f : \{-1,1\}^n \to [-1,1]$ with $\max_i \operatorname{Inf}_i(f)$ small.
2. Apply a small amount of noise / smoothing (and optionally truncate the Fourier expansion) so that we can view the resulting object as a *low-degree* multilinear polynomial P.
3. Use invariance to transfer statements about $P(X)$ (Boolean input) to $P(G)$ (Gaussian input).
4. Clamp to the range $[-1,1]$ (if necessary) to obtain a bounded function to which Borell's Gaussian theorem applies.
5. Translate the Gaussian bound back to the Boolean cube, losing only an ε error.

A key technical subtlety is that while f is bounded on $\{-1,1\}^n$, its multilinear extension $P : \mathbb{R}^n \to \mathbb{R}$ need *not* be bounded on Gaussian inputs. To address this one considers the *clipped* function

$$\operatorname{clip}(t) = \begin{cases} -1, & t < -1, \\ t, & -1 \le t \le 1, \\ 1, & t > 1, \end{cases} \qquad \text{and sets} \qquad \widetilde{P}(z) := \operatorname{clip}(P(z)).$$

By anti-concentration and hypercontractive estimates, $P(G)$ is very unlikely to land far outside $[-1,1]$ when influences are small and P has low degree, so $\widetilde{P}(G)$ and $P(G)$ have essentially the same noise stability. The invariance principle (applied to a smooth approximation of clip) then implies that the same holds on the Boolean cube. Finally Borell's theorem applies to $\widetilde{P}(G)$, giving the half-space benchmark and hence MIS.

17.5 Proof sketch of the invariance principle

We sketch the standard proof strategy (Lindeberg replacement), as in the handwritten notes.

Hybrid inputs. Let $X = (X_1, \dots, X_n)$ be uniform on $\{-1,1\}^n$ and let $G = (G_1, \dots, G_n) \sim \mathcal{N}(0, I_n)$. For $i = 0, 1, \dots, n$ define a *hybrid* vector

$$Y^{(i)} := (G_1, \dots, G_i, X_{i+1}, \dots, X_n).$$

Thus $Y^{(0)} = X$ and $Y^{(n)} = G$.

Fix a multilinear polynomial P and a smooth test function φ, and abbreviate $F(z) := \varphi(P(z))$. By telescoping,

$$|\mathbb{E}[F(X)] - \mathbb{E}[F(G)]| \le \sum_{i=1}^{n} \left| \mathbb{E}[F(Y^{(i)})] - \mathbb{E}[F(Y^{(i-1)})] \right| =: \sum_{i=1}^{n} \Delta_i.$$

Taylor expansion and moment matching. Fix i and think of the ith coordinate as the only random variable, with all others frozen. Because P is multilinear, we can decompose

$$P(z) = z_i\, L_i(z_{\neq i}) + R_i(z_{\neq i}),$$

where L_i and R_i do not depend on z_i. Now perform a Taylor expansion of φ around the point R_i:

$$\begin{aligned}\varphi(P(z)) &= \varphi(R_i + z_i L_i) \\ &= \varphi(R_i) + \varphi'(R_i)\, z_i L_i + \frac{\varphi''(R_i)}{2}\, z_i^2 L_i^2 + \frac{\varphi^{(3)}(R_i)}{6}\, z_i^3 L_i^3 + \frac{\varphi^{(4)}(\xi)}{4!}\, z_i^4 L_i^4,\end{aligned}$$

for some ξ between R_i and $R_i + z_i L_i$.

Taking expectations over z_i equal to X_i or G_i, the first three terms match because X_i and G_i have the same first three moments:

$$\mathbb{E}[X_i] = \mathbb{E}[G_i] = 0, \qquad \mathbb{E}[X_i^2] = \mathbb{E}[G_i^2] = 1, \qquad \mathbb{E}[X_i^3] = \mathbb{E}[G_i^3] = 0.$$

Thus the difference is controlled by the fourth-order remainder:

$$\Delta_i \;\lesssim\; \left\| \varphi^{(4)} \right\|_\infty \cdot \left(\frac{\mathbb{E}[X_i^4] + \mathbb{E}[G_i^4]}{4!} \right) \cdot \mathbb{E}[L_i^4].$$

Since $\mathbb{E}[X_i^4] = 1$ and $\mathbb{E}[G_i^4] = 3$, we may absorb the factor into a universal constant and write

$$\Delta_i \;\le\; \left\| \varphi^{(4)} \right\|_\infty \cdot C_0 \cdot \mathbb{E}[L_i^4].$$

Bounding $\mathbb{E}[L_i^4]$ by hypercontractivity. To control the fourth moment of L_i, we use a hypercontractive estimate (Bonami's lemma).

Lemma 17.14 (Bonami's lemma (a useful 2–4 hypercontractive form)). *Let $Q : \{-1,1\}^m \to \mathbb{R}$ be a multilinear polynomial of degree at most d. Then*

$$\mathbb{E}[Q(X)^4] \;\le\; 9^d \cdot \left(\mathbb{E}[Q(X)^2] \right)^2.$$

Applying this to L_i (whose degree is at most $D-1$) gives

$$\mathbb{E}[L_i^4] \;\le\; 9^{D-1}\cdot\left(\mathbb{E}[L_i^2]\right)^2.$$

Moreover, $\mathbb{E}[L_i^2]$ is exactly the influence of coordinate i.

Claim 17.15. *For a multilinear polynomial P,*

$$\mathbb{E}[L_i(X)^2] = \mathrm{Inf}_i(P).$$

Proof. Write the Fourier expansion $P(x) = \sum_S \widehat{P}(S)\chi_S(x)$. Then

$$L_i(x) = \sum_{S\ni i} \widehat{P}(S)\chi_{S\setminus\{i\}}(x),$$

so by orthonormality of characters,

$$\mathbb{E}[L_i(X)^2] = \sum_{S\ni i} \widehat{P}(S)^2 = \mathrm{Inf}_i(P).$$

□

Combining the last two displays,

$$\Delta_i \;\le\; \left\|\varphi^{(4)}\right\|_\infty \cdot C_1\cdot 9^D\cdot \mathrm{Inf}_i(P)^2,$$

for a universal constant C_1.

Summing over i. Summing over all coordinates,

$$|\mathbb{E}[\varphi(P(X))] - \mathbb{E}[\varphi(P(G))]| \le \sum_{i=1}^n \Delta_i \le \left\|\varphi^{(4)}\right\|_\infty \cdot C_1\cdot 9^D\cdot \sum_{i=1}^n \mathrm{Inf}_i(P)^2.$$

If $\max_i \mathrm{Inf}_i(P)\le\delta$, then

$$\sum_{i=1}^n \mathrm{Inf}_i(P)^2 \;\le\; \delta\sum_{i=1}^n \mathrm{Inf}_i(P).$$

Finally, for a degree-D multilinear polynomial one has $\sum_i \mathrm{Inf}_i(P) = \sum_S |S|\widehat{P}(S)^2 \le D\sum_S \widehat{P}(S)^2$, so after normalizing $\mathbb{E}[P(X)^2]$ (or working with bounded P) we obtain the advertised bound of order $D\cdot 9^D\cdot\delta\cdot\left\|\varphi^{(4)}\right\|_\infty$.

This completes the proof sketch of the invariance principle as presented in the notes.