# Computer Science 294 Lecture 17 Notes

Daniel Raban

March 14, 2023

# 1 Bonami's Lemma and Hypercontractivity

## 1.1 $B$-reasonable random variables and Bonami's lemma

**Definition 1.1.** For a real $B \geq 1$, a random variable $X$ is **$B$-reasonable** if

$$\mathbb{E}[X^4] \leq B(\mathbb{E}[X^2])^2.$$

This property is scaling invariant: If $X$ is $B$-reasonable, then $cX$ is also $B$-reasonable for $B \in \mathbb{R}$.

**Example 1.1.** The constant random variable $X = 1$ is 1-reasonable.

**Example 1.2.** $X \sim \{\pm 1\}$ is 1-reasonable.

**Example 1.3.** A standard Gaussian random variable $Z \sim N(0, 1)$ is 3-reasonable.

**Example 1.4.** The uniform random variable $Y \sim U[-1, 1]$ is $\frac{9}{5}$-reasonable.

**Example 1.5.** Let

$$X = \begin{cases} 1 & \text{with probability } 1/2^n \\ 0 & \text{with probability } 1 - 1/2^n. \end{cases}$$

Then $\mathbb{E}[X^4] = 1/2^n$ and $\mathbb{E}[X^2] = 1/2^n$, so $B = 2^n$. We think of this random variable as not so reasonable because with a small probability, it can be 1, which "surprises" us.

**Proposition 1.1** (Concentration)**.** *Let $X \geq 0$ a.s. If $X$ is $B$-reasonable, then*

$$\mathbb{P}(X \geq t\|X\|_2) \leq \frac{B}{t^4}.$$

Here, $\|X\|_2 := \sqrt{\mathbb{E}[X^2]}$. We can restate the definition of a $B$-reasonable random variable as

$$\|X\|_4 \leq B^{1/4}\|X\|_2.$$

*Proof.* Use Markov's inequality. $\square$

**Remark 1.1.** Compare this concentration result with Chebyshev's inequality, which gives a rate of decay of $1/t^2$.

**Proposition 1.2** (Anti-concentration)**.** *Let $X \geq 0$ a.s. If $X$ is $B$-reasonable, then for $0 \leq t \leq 1$,*

$$\mathbb{P}(X \geq t\|X\|_2) \geq \frac{(1-t^2)^2}{B}.$$

We will not prove this, but you can prove it as an exercise or find the proof in the textbook.

**Lemma 1.1** (Bonami)**.** *Let $f : \{\pm 1\}^n \to \mathbb{R}$ have $\deg f \leq k$, and let $X \sim \{\pm 1\}^n$. Then $f(X)$ is $9^k$-reasonable. That is,*

$$\mathbb{E}_{X \sim \{\pm 1\}^n}[f(X)^4] \leq 9^k (\mathbb{E}_{X \sim \{\pm 1\}^n}[f(X)^2])^2.$$

We can also think of this as

$$\|f\|_4 \leq (\sqrt{3})^k \|f\|_2.$$

*Proof of Bonami's lemma.* Recall the derivative

$$D_i f(x) = \frac{f(x^{i \mapsto +1}) - f(x^{i \mapsto -1})}{2}.$$

We can similarly define the operator

$$E_i f(x) = \frac{f(x^{i \mapsto +1}) + f(x^{i \mapsto -1})}{2}.$$

These two give a representation of $f$:

$$f(x) = X_i \cdot D_i f(x) + E_i f(x).$$

Note that $D_i f$ and $E_i f$ are functions that don't depend on $x_i$. In particular, $D_i f$ has degree $\leq k-1$, and $E_i f$ has degree $\leq k$.

Now we proceed by induction on $n$ and $k$ (for each $k$, we do induction on $n$, then overall induct over $k$). For $n = 0$, a degree 0 function is a constant, so this inequality holds. If $n \geq 1$, write

$$f(x) = X_n \cdot D_n f(x) = E_n f(x),$$

and define the random variables

$$d = D_n f(X), \qquad e = E_n f(X), \qquad \text{where } X_1, \ldots, X_{n-1} \sim \{\pm 1\}.$$

Then

$$\mathbb{E}_{X\sim\{\pm1\}^n}[f(X)^4] = \mathbb{E}[(X_n d + e)^4],$$

which we want to compare to

$$9^k(\mathbb{E}[f(X)^2])^2 = 9^k(\mathbb{E}[X_n d + w)^2])^2.$$

To compare these, use the binomial theorem and the linearity of expectation:

$$\begin{aligned}\mathbb{E}[(X_n d + e)^4] &= \mathbb{E}[X_n^4 d^4] + 4\cancel{\mathbb{E}[X_n^3 d^3 e]} + 6\,\mathbb{E}[X_n^2 d^2 e^2] + 4\cancel{\mathbb{E}[X_n d e^3]} + \mathbb{E}[e^4] \\ &= \mathbb{E}[d^4] + 6\,\mathbb{E}[d^2 e^2] + \mathbb{E}[e^4].\end{aligned}$$

On the other side, we have

$$\begin{aligned}9^k(\mathbb{E}[(X_n d + w)^2])^2 &= 9^k(\mathbb{E}[X_n^2 d^2] + 2\cancel{\mathbb{E}[X_n d e]} + \mathbb{E}[e^2])^2 \\ &= 9^k((\mathbb{E}[d^2])^2 + 2\,\mathbb{E}[d^2]\,\mathbb{E}[e^2] + (\mathbb{E}[e^2])^2)\end{aligned}$$

By the inductive hypothesis, we know that $\mathbb{E}[d^4] \le 9^{k-1}(\mathbb{E}[d^2])^2$ and $\mathbb{E}[d^4] \le 9^k(\mathbb{E}[e^2])^2$. By Cauchy-Schwarz,

$$\begin{aligned}6\,\mathbb{E}[d^2e^2] &\le 6\sqrt{\mathbb{E}[d^4]\,\mathbb{E}[e^4]} \\ &\le 6\sqrt{9^{k-1}}\,\mathbb{E}[d^2]\sqrt{9^k}\,\mathbb{E}[e^2] \\ &= 2\cdot 9^k\,\mathbb{E}[d^2]\,\mathbb{E}[e^2].\end{aligned}$$

Now compare the two sides term by term. □

Recall the noise operator $T_\rho f(x) = \mathbb{E}_{Y\ \rho\text{-corr. with } x}[f(Y)]$. We know that $\widehat{T_\rho f}(S) = \rho^{|S|/2}\widehat{f}(S)$.

**Corollary 1.1.** *Let $f : \{\pm1\}^n \to \mathbb{R}$, and let $f^{=k} = \sum_{|S|=k}\widehat{f}(S)^2\chi_S$. Then*

$$\|T_{1/3}f^{=k}\|_4 \le \|f^{=k}\|_2.$$

*Proof.* Using Bonami's lemma,

$$\|(\tfrac{1}{\sqrt{3}})^k f^{=k}\|_4 = (\tfrac{1}{\sqrt{3}})^k\|f^{=k}{}_4 \le \|f^{=k}\|_2.$$ □

## 1.2 Hypercontractivity and its consequences

**Theorem 1.1** ((4, 2)-hypercontractivity)**.** *For all $f : \{\pm1\}^n \to \mathbb{R}$,*

$$\|T_{1/\sqrt{3}}f\|_4 \le \|f\|_2.$$

More generally, we can say $\|T_{1/\sqrt{q-1}}f\|_q \leq \|f\|_2$ for all $q \geq 2$.

*Proof.* The proof is similar to the proof of Bonami's lemma. □

Bonami's lemma tells you how reasonable low-degree functions are. On the other hand, hypercontractivity tells you how smooth the function becomes once you apply the noise operator.

**Theorem 1.2** ((2, 4/3)-hypercontractivity)**.** *For all* $f : \{\pm 1\}^n \to \mathbb{R}$*,*

$$\|T_{1/\sqrt{3}}f\|_2 \leq \|f\|_{4/3}.$$

*Proof.*

$$\begin{aligned}
\|T_{1/\sqrt{3}}f\|_2^2 &= \mathbb{E}[T_{1/\sqrt{3}}f(X)^2] \\
&= \langle T_{1/\sqrt{3}}f(X)^2, T_{1/\sqrt{3}}f(X)^2\rangle \\
&= \sum_S \left(\frac{1}{\sqrt{3}}\right)^{|S|} \widehat{f}(S) \left(\frac{1}{\sqrt{3}}\right)^{|S|} \widehat{f}(S) \\
&= \sum_S \left(\frac{1}{3}\right)^{|S|} \widehat{f}(S)^2 \\
&= \langle f, T_{1/3}f\rangle
\end{aligned}$$

Using Hölder's inequality with $p = 4/3$ and $q = 4$,

$$\leq \|f\|_{4/3}\|T_{1/3}f\|_4$$

Using the (4, 2)-Hypercontractivity with $T_{1/\sqrt{3}}(T_{1/\sqrt{3}}f)$,

$$\leq \|f\|_{4/3}\|T_{1/\sqrt{3}}f\|_2.$$

Now divide both sides by $\|T_{1/\sqrt{3}}f\|_2$. □

**Corollary 1.2.** *For all* $f : \{\pm 1\}^n \to \mathbb{R}$*,*

$$\mathrm{Stab}_{1/3}(f) \leq \|f\|_{4/3}^2.$$

*Proof.*

$$\begin{aligned}
\mathrm{Stab}_{1/3}(f) &\leq \langle T_{1/\sqrt{3}}f, T_{1/\sqrt{3}}f\rangle \\
&\leq \|T_{1/\sqrt{3}}f\|_2^2 \\
&\leq \|f\|_{4/3}^2.
\end{aligned}$$

□

**Corollary 1.3** (Small-set expansion of noisy hypercube)**.** *Let $f : \{\pm 1\}^n \to \{0,1\}$ and let $\alpha = \mathbb{E}[f] = \mathbb{P}_{X \sim \{\pm 1\}^n}(f(X) = 1)$. Then*

$$\mathrm{Stab}_{1/3}(f) \leq \left( \frac{(\mathbb{E}[|f(X)|^{4/3}])^{3/4}}{\alpha} \right)^2 = \alpha^{3/2}.$$

**Remark 1.2.** This has a combinatorial interpretation:

$$\begin{aligned} \mathrm{Stab}_{1/3}(f) &= \mathbb{E}_{(X,Y)\ 1/3\text{-corr.}}[f(X)f(Y)] \\ &= \mathbb{P}(f(X) = 1)\mathbb{P}(f(Y) = 1 \mid f(X) = 1) \\ &= \alpha \mathbb{P}(f(Y) = 1 \mid f(X) = 1), \end{aligned}$$

So this says that

$$\mathbb{P}(f(Y) = 1 \mid f(X) = 1) \leq \alpha^{1/2}.$$

Thinking of this in terms of a random walk on the hypercube, if we take $f(x) = \mathbb{1}_{\{x \in A\}}$, we get behavior of the form

$$\mathbb{P}(Y \notin A \mid X \in A) \geq 1 - \alpha^{1/2}.$$

Here is a key corollary:

**Corollary 1.4.** *Let $f : \{\pm 1\}^n \to \{\pm 1\}$, and let $g = D_i f : \{\pm 1\}^n \to \{+1, 0, -1\}$. Define the $1/3$-**influence** as*

$$\mathrm{Inf}_i^{(1/3)}(f) := \mathrm{Stab}_{1/3}(D_i f) = \sum_S \widehat{f}(S)^2 \left( \frac{1}{3} \right)^{|S|-1}.$$

*Then*

$$\mathrm{Inf}_i^{(1/3)}(f) \leq (\mathrm{Inf}_i(f))^{3/2}.$$

*Proof.* Since $D_i f$ takes $\pm 1$-values, we have

$$\begin{aligned} \mathrm{Stab}_{1/3}(D_i f) &\leq \|D_i f\|_{4/3}^2 \\ &= ((\mathbb{E}[|D_i f(X)|]^{4/3})^{3/4})^2 \\ &= ((\underbrace{\mathbb{E}[|D_i f(X)|]^2}_{\mathrm{Inf}_i(f)})^{3/4})^2 \\ &= (\mathrm{Inf}_i(f))^{3/2}. \end{aligned}$$ □

Next time, we will see an application of hypercontractivity to proving the KKL theorem:

**Theorem 1.3** (Kahn-Kalai-Linial)**.** *For* $f : \{\pm 1\}^n \to \{\pm 1\}$,

$$\textit{Max Inf}(f) \geq \Omega(\mathrm{Var}(f)\frac{\log n}{n}).$$

**Remark 1.3.** Compare this to the Poincaré inequality, which gives

$$\text{Max Inf}(f) \geq \frac{\mathrm{Var}(f)}{n}.$$

We will also see Friedgut's theorem and the FKN theorem:

**Theorem 1.4** (Friedgut)**.** *Any function* $f : \{\pm 1\}^n \to \{\pm 1\}$ *is* $\varepsilon$*-close to a* $2^{O(\mathbb{I}(f))/\varepsilon}$*-junta.*

**Theorem 1.5** (Friedgut-Kalai-Naor)**.** *If* $f : \{\pm 1\}^n \to \{\pm 1\}$ *has* $W^{\leq 1}(f) \geq 1 - \varepsilon$ *(that is,* $\sum_{|S|\leq 1} \widehat{f}(S)^2 \geq 1 - \varepsilon$*), then* $f$ *is* $O(\varepsilon)$*-close to a dictator function.*