Motivation

Behaviour of a function of independent random variables $\xi_1, \ldots, \xi_n$:

$$X = F(\xi_1, \ldots, \xi_n)$$

- the random variable $X$ often counts certain objects or events

Sharp concentration: $X \approx \mathbb{E}X$

In applications we usually aim at estimates of form

$$\mathbb{P}(X \not\in (1 \pm \varepsilon)\mathbb{E}X) \leq N^{-\omega(1)}$$

- Replacing $N^{-\omega(1)}$ with $o(1)$ is frequently not good enough

Topic of his talk

*Easy-to-check* conditions which guarantee concentration
Toy-Example: Sums of iid Indicators

Chernoff–Bernstein type inequality (1952 and 1924)

Let $X = (X_1, \ldots, X_N)$ be independent 0/1 variables: $\mathbb{P}(X_i = 1) = 1/2$. For $f(X) = \sum_{1 \leq i \leq N} X_i$ we have

$$\mathbb{P}(|f(X) - \mathbb{E}f(X)| \geq t) \leq 2e^{-t^2/N}$$

Concentration follows:

- $|X - \mathbb{E}X| \leq N^{1/2 + o(1)}$ with probability $1 - N^{-\omega(1)}$

Setting of this talk

Similar result when $f(X)$ is a more complicated function of the $X_i$
Bounded differences inequality (McDiarmid, 1989)

**Lipschitz-condition:** whenever $x, \tilde{x}$ differ in one coordinate,

$$|f(x) - f(\tilde{x})| \leq c$$

If $X = (X_1, \ldots, X_N)$ are independent random variables, then

$$\mathbb{P}(|f(X) - \mathbb{E}f(X)| \geq t) \leq 2e^{-t^2/2Nc^2}$$

**Concentration follows:**

- $|f(X) - \mathbb{E}f(X)| \leq cN^{1/2+o(1)}$ with probability $1 - N^{-\omega(1)}$

**Intuitively:** this bound can’t be sharp???

- Large ‘worst case’ changes should be irrelevant
- Smaller ‘typical’ changes should matter
Typical bounded differences inequality (simplified, W.)

**Typical event** $\Gamma$:

$$\mathbb{P}(X \in \Gamma) \geq 1 - N^{-\omega(1)}$$

**Typical Lipschitz-condition**: if $x \in \Gamma$ and $\tilde{x}$ differ in one coordinate, then

$$|f(x) - f(\tilde{x})| \leq c$$

If $|f(X)| \leq N^{O(1)}$, then for independent $X = (X_1, \ldots, X_N)$ we have

$$\mathbb{P}(|f(X) - \mathbb{E}f(X)| \geq t) \leq 2e^{-t^2/3Nc^2} + N^{-\omega(1)}$$

**Punchline for concentration**:

- can replace worst case changes by typical changes
  (which makes heuristic considerations rigorous)
Typical bounded differences inequality (simplified, W.)

**Typical event** $\Gamma$: 
\[ \mathbb{P}(X \in \Gamma) \geq 1 - N^{-\omega(1)} \]

**Typical Lipschitz-condition:** if $x \in \Gamma$ and $\tilde{x}$ differ in one coordinate, 
\[ |f(x) - f(\tilde{x})| \leq c \]

If $|f(X)| \leq N^{O(1)}$, then for independent $X = (X_1, \ldots, X_N)$ we have 
\[ \mathbb{P}(|f(X) - \mathbb{E}f(X)| \geq t) \leq 2e^{-t^2/3Nc^2} + N^{-\omega(1)} \]

**Remarks:**
- $|f(X) - \mathbb{E}f(X)| \leq cN^{1/2+o(1)}$ with probability $1 - N^{-\omega(1)}$
- Matches heuristics: $c$ is now the ‘typical change’
- Conditions fairly intuitive and easy-to-check
Typical bounded differences inequality (simplified, W.)

**Typical event** $\Gamma$:

$$\mathbb{P}(X \in \Gamma) \geq 1 - N^{-\omega(1)}$$

**Typical Lipschitz-condition:** if $x \in \Gamma$ and $\tilde{x}$ differ in one coordinate,

$$|f(x) - f(\tilde{x})| \leq c$$

If $|f(X)| \leq N^{O(1)}$, then for independent $X = (X_1, \ldots, X_N)$ we have

$$\mathbb{P}(|f(X) - \mathbb{E}f(X)| \geq t) \leq 2e^{-t^2/3Nc^2} + N^{-\omega(1)}$$

‘Naive guesses’ are wrong (in general):

- $\mathbb{P}(|f(X) - \mathbb{E}f(X)| \geq t \mid X \in \Gamma) \leq e^{-\Theta(t^2/Nc^2)}$
- $\mathbb{P}(|f(X) - \mathbb{E}f(X)| \geq t) \leq e^{-\Theta(t^2/Nc^2)} + \mathbb{P}(X \not\in \Gamma)$
- $\mathbb{P}(|f(X) - \mathbb{E}f(X)| \geq t \text{ and } X \in \Gamma) \leq e^{-\Theta(t^2/Nc^2)}$
It seems to be a convenient tool (e.g., to simplify/shorten proofs)

Some applications of the typical bounded differences inequality

- **Additive combinatorics**  
  Sum-free subsets in abelian groups (Morris et. al)

- **Probabilistic combinatorics**  
  Phase transition in random graph coloring (Coja-Oghlan et. al)

- **Theoretical computer science**  
  Average case analysis of euclidean functions (de Graaf–Manthey)

- **Random graph processes**  
  $H$-free graphs (W.)

- **Applied mathematics/Electrical engineering**  
  Error-correcting codes (Häger et. al)

- ???  
  Please try your favourite problem...
Reverse $H$-free process (ends with $H$-free graph)

- Start with a complete graph $K_n$ on $n$ vertices
- In each step: a random edge is \textit{removed}, chosen uniformly from all edges that are \textit{contained} in a copy of $H$

Motivation: applications to Ramsey/Turán theory

Question of Bollobás–Erdős (1990)

What is the typical final number of edges $M = M(n, H)$?

Some answers: the final number of edges is

- Makai: whp $M \sim c_H n^{2-1/d_2(H)}$ for \textit{strictly} 2-balanced $H$
- Warnke: whp $M \sim \mathbb{E}M = \Theta(n^{2-1/d_2(H)})$ for 2-balanced $H$
Results for the Bollobás–Erdős Question

Reverse $H$-free process: the final number of edges is

- Makai: whp $M \sim c_H n^{2-1/d_2(H)}$ for strictly 2-balanced $H$
- Warnke: whp $M \sim \mathbb{E} M = \Theta(n^{2-1/d_2(H)})$ for 2-balanced $H$

- Surprise: can analyze process without differential equation method!

Proof approaches

- Makai: delicate first and second moment arguments
  (using FKG, Janson+Suen inequalities to evaluate $\mathbb{E} M^2$)
- Warnke: using TBD-inequality it is enough to calculate $\mathbb{E} M$
  (we can routinely ‘override’ the weak dependencies)
Reverse $H$-free process (alternative definition)

Order edges of complete graph $K_n$ uniformly at random ($e_1, e_2, \ldots$). Start with complete graph $K_n$ and process edges sequentially ($e_{n \choose 2}, \ldots$): remove edge if and only if it currently lies in a copy of $H$.

Key observation (due to Makai + Erdős–Suen–Winkler)

The decision whether $e_j$ is removed depends only on $(e_i)_{1 \leq i \leq j}$.

- Proof sketch: if $e_j$ lies in a copy of $H$ that contains edges $e_i$ with $i > j$, then one of these would have been removed by the process.

Surprising consequence

$e_j$ in final graph iff it closes no copy of $H$ together with $(e_i)_{1 \leq i < j}$.

- Note: $\{e_1, \ldots, e_m\} \equiv G_{n,m}$, i.e., the uniform random graph.
**Small typical changes (2/2)**

*Sketch of the argument for $H = K_3$ (triangle)*

\[ G_{n,m} \equiv \{e_1, \ldots, e_m\} \]

$e_j$ in final graph iff it closes *no* copy of $K_3$ together with \( (e_i)_{1 \leq i < j} \)

**Standard facts for $G_{n,m^*}$ with $m^* = n^{3/2}(\log n)^2$**

- Wvhp every edge of $G_{n,m^*}$ lies in at least one copy of $K_3$
- Wvhp every pair of vertices has codegree at most $\leq (\log n)^5$

**Simple proof: concentration of the final number of edges**

- Enough to study \( (e_i)_{1 \leq i \leq m^*} \), i.e., first $m^*$ edges
- **Small typical changes**: each edge influences $O((\log n)^5)$ other edges
- Typical bounded differences inequality *routinely* shows concentration
  (it also applies to $G_{n,m}$ or random permutations)
Typical bounded differences inequality (punchline)

For establishing concentration via the bounded-differences approach, we can often replace the worst case changes by the typical changes.

Remarks:
- Typical changes coincide with heuristics (whether concentration holds)
- Conditions fairly intuitive and easy-to-check
- Paper contains more power/flexible version of the inequality

Open Problem

More applications?