Recent progress on the Lova’sz Local Lemma

Aravind Srinivasan
University of Maryland, College Park
Roadmap

• Moser-Tardos, how far their ideas can be pushed
• Main message: *MT is of indep. interest, not only as an algorithmic version of the LLL*
• Proof sketch for MT (hopefully more intuitive)
• Beyond MT
• Open questions
Moser-Tardos, and how far their ideas can go
[Joint work with Bernhard Haeupler & Barna Saha, JACM ‘11]

• The Lova’sz Local Lemma (Erdos & Lovasz, 1975): a major probabilistic tool.

• Numerous app.s with only known proof via the LLL (graph coloring, scheduling, packet routing in wired/wireless networks, ...)

• Moser & Tardos: breakthrough on the algorithmic aspect.
Example: k-SAT

- Boolean variables \( x_1, x_2, \ldots, x_n \), say \( k = 4 \)
- “Literal”: can be \( x_i \) or \( \neg x_i \) (\( = \) not(\( x_i \)))
- \( m \) clauses, each an OR of exactly \( k \) literals: e.g.,
  \((x_1 \ OR \ \neg x_4 \ OR \ \neg x_6 \ OR \ x_7), \ (x_2 \ OR \ x_4 \ OR \ \neg x_8 \ OR \ x_{11}), \ldots \)

- SAT: Does there exist a satisfying assignment?
- MAX-SAT: Maximize #satisfied clauses

- Assign each \( x_i \) randomly: each clause fails with probability \( 2^{-k} \).
- \( E[\text{#failed clauses}] = m \ 2^{-k} \); useful when \( k > \log_2 m \). Small \( k \)?
**Independent** random variables \(X_1, X_2, \ldots, X_n\)

“Bad” events \(E_1, \ldots, E_m\): each a function of some subset of the \(X_j\)’s

Example (think of 3-SAT):

\[
E_1 = f_1(X_1, X_5, X_8) \\
E_2 = f_2(X_2, X_5, X_{10}) \\
E_3 = f_3(X_1, X_8, X_{10}) \\
E_4 = f_4(X_2, X_6, X_{11}) \\
\ldots
\]

LLL [Symmetric Vers.]: If all \(\Pr[E_i] \leq p\) and each \(E_i\) “depends” on at most \(d\) others, then \(e^p (d + 1) \leq 1\) implies \(\Pr[\text{all } E_i \text{ avoided}] > 0\).

**k-SAT**: Give random assignment to the variables; \(p = 2^{-k}\).
Satisfiable if each clause overlaps \(\leq d = 2^k/e - 1\) others!
Algorithm to find a good assignment?
Moser-Tardos

Algorithm:
Sample the variables from their distributions

While some $E_i$ is true:
  pick any such $E_i$ and resample its variables

MT’s Theorem: Quick convergence! At most $O(m)$ iterations in expectation.

Issues:
• m too large (e.g., exponential in $n = \text{“output size”}$)
• what if picking a true $E_i$ is (NP-)hard?

Both can occur in applications: e.g., graph coloring, max-min fair allocation (Feige’s non-constructive $O(1)$-approx.)
[HSS]: first result

• Both issues can be bypassed by allowing a small slack
• LLL-condition [General asymmetric version]: suffices to have $0 < x_i < 1$ for each $E_i$ such that

$$\Pr[E_i] \leq x_i \prod_{j \sim i} (1 - x_j)$$

**Theorem:** Suppose $\Pr[E_i]^{1-\varepsilon} \leq x_i \prod_{j \sim i} (1 - x_j)$ for all $i$. Then, suffices to run [MT] on $n^{c/\varepsilon}$-sized “core” subset of the bad events. [Core checkable in all known applications; events there depend on $O(\log n)$ variables.]

All known direct apps (except for the “lopsided” LLL) made algorithmic – with the small added slack of $\varepsilon$. 
New Theorem: For any event $A$ that is a function of the $X_j$’s, if $D =$ output distr. of [MT],

$$\Pr_D[A] \leq \Pr[A] / \prod_{j \sim A} (1 - x_j)$$

Approach: witness trees.

Corollary: Applications that allow a “few” bad events (e.g., MAX SAT): if dep. $d = \alpha * (1/\epsilon p)$, then approx. $(e \ln(\alpha) / \alpha) * mp$ violated events.

New even non-constructively; more in [Harris-S.]
[MT] Proof Idea, I

• Find reason for large # of resamplings
• Suppose $E_i$ was the $t^{th}$ resampling. Go back in time, construct *witness tree* with root $E_i$.

Time going back starting from $t$

New node $E_o$ at time $t' < t$: Make child of *deepest* neighbor (if one exists)
[MT] Proof Idea, II

- Key theorem: Fix $T$.
  $\Pr[T$ occurs as a w.t.$] = \text{weight}(T) = \prod_{E \in T} \Pr[E].$

  - Values of $X_1$: * * * * * ...
  - Values of $X_2$: * * * * * ...
  - Values of $X_3$: * * * * * ...
  - Values of $X_4$: * * * * * ...
  - Values of $X_5$: * * * * * ...

  Each level is an indep. set.
  BFS bottom-up!
  Holds for $T$ becoming a w.t. *anytime*. 
[MT] Proof Idea, III: *Simplified*

- \(#\text{resamp. of } E_i = \#\text{w.t.s with root } E_i\)
- Let \(\mu_i = \text{total weight of trees with root } E_i\)
- \(E[\#\text{resamp. of } E_i] = \mu_i\)

- MT: assume LLL hypothesis, derive \(\mu_i\)
- Can reverse this: \(\mu_i \text{ exist } \rightarrow \text{LLL hypothesis!}\)
Idea: the $\mu_i$ ‘s satisfy a simple *recurrence* – children of any node in the w.t., are indep. (hence *no repeats*)

$$E_i$$

Children of root are $E_i$ or $E_i$’s neighbors; no repeats

$$\mu_i = \Pr[E_i] \cdot \prod_{j: (j = i) \text{ or } (j \sim i)} (1 + \mu_j)$$

Existence of finite soln. $\{\mu_i\}$ does not imply convergence?
- Replace “=“ by “≥“
- Ind. on $h$: weight of trees with height $\leq h$, is $\leq \mu_i$

Set $\mu_i = x_i / (1 - x_i)$, $0 < x_i < 1$. Recover $\Pr[E_i] \leq x_i \prod_{j \sim i} (1 - x_j)$
Further improvements: Pegden and others

- Recall: children of each node in w.t. are $indep.$
- In particular, if node $u$ labeled $E_i$ has a child labeled $E_i$, it can have no other children
- Recover “$e^p D \leq 1$ (for $D \geq 1$)” [Shearer]
- Further improvements for k-SAT for small $k$
Beyond MT: Constraint Satisfaction
[Joint work with David Harris, STOC ‘13 & FOCS ‘13]

- Universal defn. with Assignment and Decreasing constraints:

- CSPs with n variables; i\textsuperscript{th} var. takes values in finite set A\textsubscript{i}:
  Assignment constr: for all i, \( \sum_{j \in A_i} x_{i,j} = 1 \) (\( x_{i,j} \in \{0,1\} \))

- Decreasing Boolean constraints: Let V = \{all var.s x_{i,j}\}. For k = 1, 2, ..., m, there is a bad event E\textsubscript{k}: “a given Boolean function B\textsubscript{k} that is an increasing function of some subset S\textsubscript{k} of V, holds”
Example: graph transversals

- Introduced by Bollobas, Erdos & Szemeredi; many applications in graph coloring and partitioning
- Graph $G = (V,E)$, max. degree $\Delta$.
- Given a partition of $V$, want a transversal (one vertex from each block – assignment constraint), with good properties – e.g., no clique of size $\geq s$ (decreasing constraint)

Min. size $b$ of each block that suffices (for all possible partitions)
“$s = 2$” well-studied: $b = 2 \Delta$ [Haxell], not yet algorithmic
Lower bound: $\left( \frac{s}{(s-1)^2} \right) \Delta$ [Szabo-Tardos], conjectured optimal
Upper bound: $2 \times \text{floor}(\Delta / (s-1))$ [Loh-Sudakov]
Standard rounding

• Given a fractional assignment $x$ for all $i$, $\sum_{j \in A_i} x_{i,j} = 1$, randomly choose exactly one $j$ for each $i$ (indep.ly for different $i$), with $\Pr[Y_{i,j} = 1] = x_{i,j}$.

• Union-bound-based analysis: want $\Pr[B_k = 1] \ll 1/m$, often too much to ask.

• Standard LLL-based analysis: even if no variable appears in many constraints, correlations among the $B_k$ can be too large.
Main idea: *partial* resampling

- Recall: each bad $E_i$ det. by some subset $S_i$ of the var.s

- Why sample *all of* $S_i$?

- Construct distribution $D_i$ on *subsets* of $S_i$; need to resample for $E_i \rightarrow$ sample subset from $D_i$, *only* resample subset

- Example: for $K_s$-free transversals, if we currently contain $C$, a $K_s$, choose a random $\sqrt{s}$-sized subset of $C$, and resample it.
Application: $K_s$-free transversals

- Recall $(s/(s-1)^2) \Delta \leq b_{opt} \leq 2 \times \text{floor}(\Delta / (s-1))$, for $K_s$-free transversals
- Here: lower-bound asymptotically true:
  \[ b_{opt} \leq \left(1/s + O(s^{-1.5})\right) \Delta \]
Application: generalized multi-dim. scheduling

- m machines, n jobs; assign each job to one machine
- D dimensions to load on each machine (e.g., time, storage, energy consumed, ...)
- Assigning job j to machine i, adds $p_{i,j,k}$ to load on dim. k on machine i
- Given $(T_1, T_2, ..., T_D)$, is there an assignment with max. load $T_d$ in each dim. d?
- Azar-Epstein’s approach: $(D + 1)$--approximation
- Here: $O(\log D / \log\log D)$--approx.
Application: packet routing

- Seminal Leighton-Maggs-Rao result: given k (source-dest) pairs each with a routing path, schedule packets (≤ 1 packet moves on an edge in a step) to minimize delivery time.

- Let $C = \text{congestion}$ (max. #given paths that use an edge), and $D = \text{dilation}$ (max. length of any given routing path).

- Easy: $\text{OPT} \geq C$, $\text{OPT} \geq D$; so $\text{OPT} \geq (C + D)/2$.

- [LMR ‘88]: $\text{OPT} \leq O(C + D)$ always (iterated LLL-based proof with much impact); constant very large, constructive.

- [Peis-Wiese ‘11]: $\text{OPT} \leq 23.4(C+D)$; constructive LLL-based.

- Here (incl. symmetry-group ideas): $\text{OPT} \leq 5.4(C + D)$, constr.
Open questions and Recent Progress

• Further understanding of [MT]?
• LMR constant?
• Better (algorithmic) understanding of Shearer’s characterization of the LLL?
• Finer classification into positive/negative correlation: the lopsided LLL (joint work with David Harris, SODA 2014)
  – random permutations, appear much more complex
  – apps. in discrete math (Latin transversals, rainbow cycles, ...)