More on fluctuations in first passage percolation

Thursday, 27 May 2021 13:03

Detour to influences in Boolean fcns. and Talagrand's inequality (following Benjamini–Kalai–Schramm 2003)

Hypercube: $\{0,1\}^n$
Equipped with the uniform measure $\mathbb{P}$.
Random variables: $f:\{0,1\}^n \to \mathbb{R}$.

Influence: For $1 \le j \le n$, let $\sigma_j f:\{0,1\}^n \to \mathbb{R}$

$$(\sigma_j f)(x) := f(x_1, \ldots, x_{j-1}, 1-x_j, x_{j+1}, \ldots, x_n)$$

flip j'th coordinate operation

$$(\rho_j f)(x) := \frac{f(x) - (\sigma_j f)(x)}{2}$$

The function measures the change made by the flip. The norms of $\rho_j f$ can be thought of as the influence of the j'th coordinate.

Norms: $\|f\|_p := (\mathbb{E}(|f|^p))^{1/p} = \left(\frac{1}{2^n}\sum_{x\in\{0,1\}^n} |f(x)|^p\right)^{1/p}$.

Theorem (Talagrand 1994):

$$\mathrm{Var}(f) \le C \sum_{j=1}^n \frac{\|\rho_j f\|_2^2}{1 + \log(\|\rho_j f\|_2 / \|\rho_j f\|_1)}$$

Without the denominator, this is in the same vein as the Efron–Stein inequality.
Get a significant improvement if $\|\rho_j f\|_2 >> \|\rho_j f\|_1$

This is the case, in our application, when $\rho_j f$ is only rarely non-zero.

I'll show a proof following

We show a proof following Benjamini–Kalai–Schramm (2003). The proof relies on the Bonami–Beckner inequality
1970 1975

Bonami-Beckner: Noise operator N_ε $(0 \le \varepsilon \le 1)$ takes a function $f: \{0,1\}^n \to \mathbb{R}$ and returns another fcn. which is somewhat averaged.

Let $X_1, \dots, X_n$ be Bernoulli(ε) indep

$$X_i \sim \begin{cases} 1 & \varepsilon \\ 0 & 1-\varepsilon \end{cases}.$$

$$N_\varepsilon(f)(x) = \mathbb{E}(f(x + X \bmod 2))$$

Each coord. is flipped with prob. ε and we average f on the resulting random position.

Theorem (Bonami-Beckner): For each $0<\varepsilon<1$

$$\|N_\varepsilon f\|_2 \le \|f\|_{1+\delta(\varepsilon)}$$

Hypercontractive inequality

for $\delta(\varepsilon) = (1-2\varepsilon)^2$.

Representation in a Fourier-Walsh basis:

For a subset $S \subseteq \{1, \dots, n\} =: [n]$, let $\chi_S : \{0,1\}^n \to \{-1,1\}$ be

$$\chi_S(x) = (-1)^{\sum_{i \in S} x_i}. \qquad (\chi_\emptyset \equiv 1).$$

Fourier-Walsh basis

This is an orthonormal basis for L^2 of fcns. on the hypercube with the uniform measure.

Parseval: Write $f(x) = \sum \hat{f}(S) \chi_S(x)$

or ,

Parseval: Write $f(x) = \sum_{S \subseteq [n]} \hat{f}(S) \chi_S(x)$ ← coefficients in the expansion

wrt. to the uniform measure

$$\|f\|_2^2 = \sum_{S \subseteq [n]} |\hat{f}(S)|^2.$$ ← wrt. the counting measure

Noise operator: What is the expansion of $N_\varepsilon f$?

Since $f(x) = \sum_{S \subseteq [n]} \hat{f}(S) \chi_S(x)$,

then
$$N_\varepsilon(f)(x) = \sum_{S \subseteq [n]} \hat{f}(S) \mathbb{E}(\chi_S(x + X \bmod 2))$$

$$= \sum_{S \subseteq [n]} \hat{f}(S) \mathbb{E}\left[(-1)^{\sum_{i \in S} x_i + X_i}\right] =$$

$$= \sum_{S \subseteq [n]} \hat{f}(S) \chi_S(x) \prod_{i \in S} \underbrace{\mathbb{E}\left[(-1)^{X_i}\right]}_{= 1 - \varepsilon - \varepsilon = 1 - 2\varepsilon} =$$

$$= \sum_{S \subseteq [n]} (1 - 2\varepsilon)^{|S|} \hat{f}(S) \chi_S(x).$$

Useful notation: For $-1 < p < 1$, write T_p for the operator taking f to
$$T_p(f)(x) = \sum_{S \subseteq [n]} p^{|S|} \hat{f}(S) \chi_S(x)$$

So that $T_{1-2\varepsilon} = N_\varepsilon$.

In this notation, Bonami-Beckner says that
$$\|T_p\|_2 \le \|f\|_{1+p^2}.$$

Proof of Talagrand's inequality:

Proof of Talagrand's inequality:

Write $f(x) = \sum_{S \subseteq [n]} \hat{f}(S) \chi_S(x)$.

Recall $(\rho_i f)(x) = \frac{f(x) - (\sigma_i f)(x)}{2}$ flip i'th input bit operation.

What is the Fourier-Walsh expansion of $\rho_i f$?

Simple to see that $(\sigma_i \chi_S)(x) = \begin{cases} \chi_S(x) & i \notin S \\ -\chi_S(x) & i \in S \end{cases}$

$$\Rightarrow \rho_i(f)(x) = \sum_{\substack{S \subseteq [n] \\ i \in S}} \hat{f}(S) \chi_S(x).$$

How to express the variance?

$$\mathrm{Var}(f) = \mathbb{E}(f - \underbrace{\mathbb{E}(f)}_{=\hat{f}(\emptyset)})^2 = \sum_{\emptyset \neq S \subseteq [n]} \hat{f}(S)^2.$$

Parseval

How to use the Bonami-Beckner ineq.?

$$\mathrm{Var}(f) = \sum_{\emptyset \neq S \subseteq [n]} \hat{f}(S)^2 = \sum_{i=1}^{n} \sum_{0 \neq S \subseteq [n]} \frac{\widehat{(\rho_i f)}(S)^2}{|S|} \leq$$

$$\leq 3 \sum_{i=1}^{n} \int_0^1 \| T_p \underbrace{(\rho_i f)}_{\text{Fourier-Walsh expansion} = \sum_{S \subseteq [n],\, i \in S} p^{|S|} \hat{f}(S) \chi_S(x)} \|_2^2 \, dp$$

$$\Rightarrow \| T_p(\rho_i f) \|_2^2 = \sum_{\substack{S \subseteq [n] \\ i \in S}} p^{2|S|} \hat{f}(S)^2$$

Main inequality,

Bonami-Beckner $\Rightarrow$ $\mathrm{Var}(f) \leq 3 \sum^{n} \int^{1} \| \rho_i f \|^2 \ldots dp$

Main inequality,

$j \in S$

Bonami-Beckner $\Rightarrow \mathrm{Var}(F) \leq 3 \sum_{j=1}^{n} \int_0^1 \|\rho_j F\|_{1+p^2}^2 \, dp.$

Algebraic manipulations: Hölder's ineq.

$$\mathbb{E}(|g|^{1+p^2}) = \mathbb{E}\left[|g|^{2p^2} \cdot |g|^{1-p^2}\right] \leq (\mathbb{E}|g|^2)^{p^2} \cdot (\mathbb{E}|g|)^{1-p^2}$$

$$= \|g\|_2^{2p^2} \cdot \|g\|_1^{1-p^2} = \|g\|_2^{1+p^2} \left(\frac{\|g\|_1}{\|g\|_2}\right)^{1-p^2}.$$

$$\Rightarrow \|g\|_{1+p^2}^2 \leq \|g\|_2^2 \cdot \left(\frac{\|g\|_1}{\|g\|_2}\right)^{\frac{2(1-p^2)}{1+p^2}}.$$

Also note that for $0 \leq x \leq 1$:

$$\int_0^1 x^{\frac{2(1-p^2)}{1+p^2}} \, dp \leq \ldots \leq \frac{1-x^{6/5}}{\mathrm{Log}(\frac{1}{x})}.$$

Putting everything together:

$$\mathrm{Var}(F) \leq 3 \sum_{j=1}^{n} \|\rho_j F\|_2^2 \cdot \frac{1-(\|\rho_j F\|_1 / \|\rho_j F\|_2)^{6/5}}{\mathrm{Log}(\|\rho_j F\|_2 / \|\rho_j F\|_1)}$$

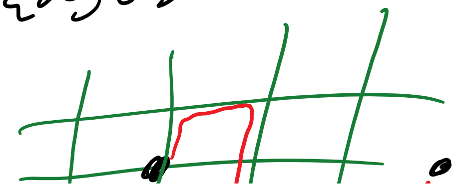
one checks

$$\leq C \sum_{j=1}^{n} \frac{\|\rho_j F\|_2^2}{1+\mathrm{Log}(\|\rho_j F\|_2 / \|\rho_j F\|_1)}$$

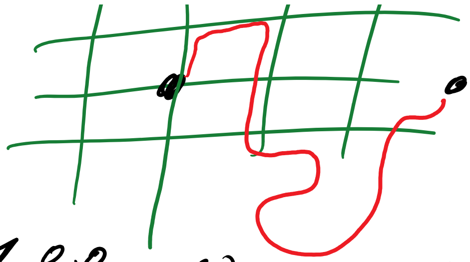
finishing the proof of Talagrand's ineq.

Application to first passage percolation

Consider FPP on $\mathbb{Z}^d$ in which the edge weights are uniform on $\{a, b\}$ for some $0 < a < b < \infty$.



For some $0<a<b<\infty$.



Assume $d \geq 2$.

Consider $T(0, Le_1)$ ($e_1 = (1,0,0,\dots,0)$, first coord. dir.)
which is the shortest passage time from 0 to Le_1.

Thm. (Benjamini–Kalai–Schramm 2003):

$$\mathrm{Var}(T(0, Le_1)) \leq C \frac{L}{\log L}.$$

(Best known upper bound)
(Believed in $d=2$ that $\mathrm{var} \approx L^{2/3}$ and even smaller in $d \geq 3$).

Note that the optimal path can use at most $\frac{b}{a} \cdot L$ edges since the straight line path is a candidate.
Thus $T(0, Le_1)$ is a fcn. of finitely many Boolean random variables. So Talagrand's ineq. applies.

For brevity, write $f = T(0, Le_1)$, a fcn. of (η_e), where $\eta_e \in \{a, b\}$ is the weight of the edge e.

Note $\partial_e f = \frac{f - \sigma_e f}{2}$, where $\sigma_e f$ is the passage time after flipping the weight of the edge e.

Note: if $\eta_e = a$ then $\sigma_e f \geq f$ always

Note: if $\eta_e = a$ then $\sigma_e f \geq f$ always and $\sigma_e f > f$ iff all shortest paths from 0 to Le_1 pass through e.

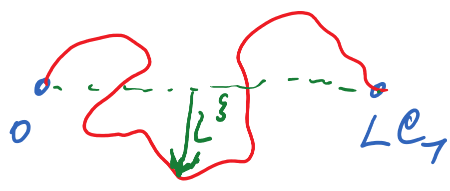
Thus $\mathbb{E}(\rho_e f)^2 = \frac{1}{2}\mathbb{E}((\rho_e f)^2 \mid \eta_e = a) =$

$$= \mathbb{P}\left(\begin{array}{l}\text{all optimal paths} \\ \text{pass through } e\end{array} \,\middle|\, \eta_e = a\right) \cdot c(a,b).$$

To improve upon Efron-Stein, need that $\dfrac{\|\rho_e f\|_2}{\|\rho_e f\|_1}$ is large.

But $$\frac{\|\rho_e f\|_2}{\|\rho_e f\|_1} = \frac{c(a,b)}{\sqrt{\mathbb{P}\left(\begin{array}{l}\text{all optimal paths} \\ \text{pass through } e\end{array} \,\middle|\, \eta_e = a\right)}}$$

This probability is expected to be small for edges far from 0 and Le_1, probably as small as $\frac{1}{L^c}$, and this would imply the BKS upper bound on the variance, by Talagrand's ineq.



Open question: Prove such an upper bound on the probability (BKS midpoint problem).

Known: The above prob. tends to 0 with L for e far from 0 and Le_1.
(Ahlberg-Hoffman 2016)

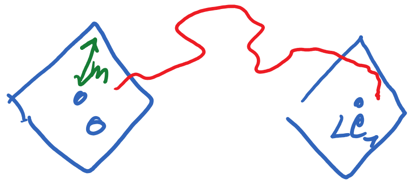
Idea to bypass the open question:

Idea to bypass the open question:

Instead of $T(0, Le_1)$ consider an averaged version:

$$S(0, Le_1) := \frac{1}{|B(0,m)|} \sum_{z \in B(0,m)} T(z, z + Le_1)$$

where $B(0,m) := \{ z \in \mathbb{Z}^d : \|z\|_1 \leq m \}$, and choosing $m := L^{1/4}$.



Certainly, $\mathbb{E}\, S(0, Le_1) = \mathbb{E}\, T(0, Le_1)$ by translation invariance.

Additionally, $|S(0, Le_1) - T(0, Le_1)| \leq \underbrace{C(a,b)}_{2b} \cdot m$

$$\Rightarrow Var(T) = \|T - \mathbb{E}T\|_2^2 \leq 2 \underbrace{\|S - \mathbb{E}S\|_2^2}_{= Var(S)} + 2(C(a,b)m)^2.$$

$\Rightarrow$ It suffices to use Talagrand's ineq. for S. The advantage is that for S, due to the averaging, one can show that $\|\partial_e S\|_2 / \|\partial_e S\|_1$ is a power of L^{-1} without much difficulty (see 50 years of FPP survey).

Final remarks: The upper bound $Var(T) \leq \frac{CL}{\log L}$ is now known for most weight dist.. Talagrand's ineq. is sometimes replaced by Falik-Smorodnitsky's ineq..

Damron-Hanson-Sosoe 2015

Detour to total variation distances

Detour to total variation distances and a lower bound on the variance

Best-known lower bound on the variance,

Thm. (Newman-Piza 1995): In $d=2$,

$$\mathrm{Var}(T(0, Le_1)) \geq c \log 2.$$

Open whether $\mathrm{Var} \xrightarrow[L\to\infty]{} \infty$ in $d \geq 3$ (open even in physics literature).

in $d=2$, expect $\mathrm{Var} \approx L^{2/3}$.

Total variation distance:

μ, ν are prob. measures on some meas. space.

$$d_{TV}(\mu, \nu) := \sup_{A \text{ event}} |\mu(A) - \nu(A)|$$

For every A, $\nu(A) - d_{TV}(\mu,\nu) \leq \mu(A) \leq \nu(A) + d_{TV}(\mu, \nu)$

When μ, ν have densities wrt. to a third measure λ (always ok with, say, $\lambda = \mu + \nu$) (usually, λ = Lebesgue or counting)

Write $\mu = f\,d\lambda$, $\nu = g\,d\lambda$.

Claim: $d_{TV}(\mu, \nu) = \frac{1}{2}\int |f-g|\,d\lambda =$

$$= \int (f-g)\mathbb{1}_{f>g}\,d\lambda = \int (g-f)\mathbb{1}_{g>f}\,d\lambda.$$

Proof: First, $\frac{1}{2}\int|f-g|\,d\lambda =$

$$= \frac{1}{2}\left[\int (f-g)\mathbb{1}_{f>g}\,d\lambda + \int (g-f)\mathbb{1}_{g>f}\,d\lambda\right] =$$

$$= \frac{1}{2}\left[\int (f-g)\mathbb{1}_{f>g}\, d\lambda + \underbrace{\int (g-f)\, d\lambda}_{=0} - \int (g-f)\mathbb{1}_{f<g}\, d\lambda\right]$$

$$= \int (f-g)\mathbb{1}_{f>g}\, d\lambda .$$

Now, for any event A,

$$\mu(A) = \int_A f\, d\lambda \le \int_A g\, d\lambda + \int_A (f-g)\mathbb{1}_{f>g}\, d\lambda \le$$

$$\le \nu(A) + \int (f-g)\mathbb{1}_{f>g}\, d\lambda$$

$$\Rightarrow d_{TV}(\mu,\nu) \le \int (f-g)\mathbb{1}_{f>g}\, d\lambda .$$

For other direction,

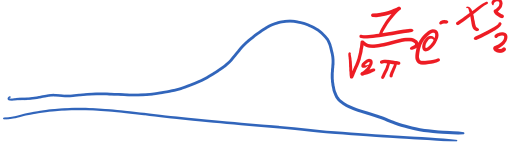
if $A = \{f > g\}$ then

$$\mu(A) = \int_A f\, d\lambda = \underbrace{\int_A g\, d\lambda}_{=\nu(A)} + \int (f-g)\mathbb{1}_{f>g} .$$

It is not always easy to give an upper bound to $d_{TV}(\mu,\nu)$. We discuss now one of the simplest cases.

Example: $X = (X_1, \ldots, X_n)$ IID, $X_i \sim N(0,1)$

and $Y = (Y_1, \ldots, Y_n)$ indep.



where $Y_i \sim N(\varepsilon_i, 1)$.

How big is $d_{TV}(\mathcal{L}(X), \mathcal{L}(Y))$?

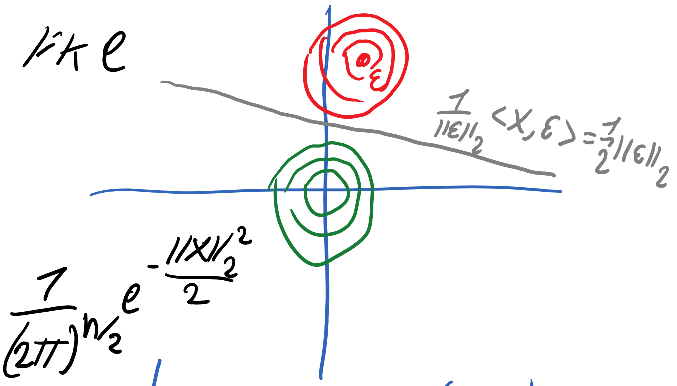
distributed as $N(0,1)$

Claim: $d_{TV}(\mathcal{L}(X), \mathcal{L}(Y)) = \mathbb{P}\left(\frac{1}{\|\varepsilon\|_2}|\langle X, \varepsilon\rangle| < \frac{1}{2}\|\varepsilon\|_2\right)$

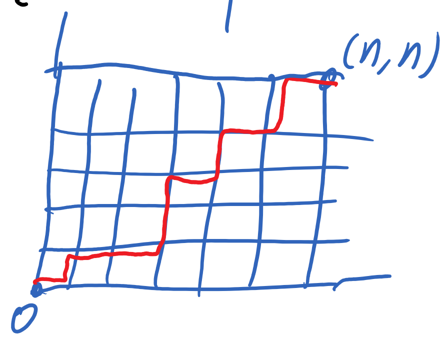
Claim: $d_{TV}(\mathcal{L}(X), \mathcal{L}(Y)) = \mathbb{P}\left(\frac{1}{\|\varepsilon\|_2}|\langle X, \varepsilon\rangle| < \frac{1}{2}\|\varepsilon\|_2\right)$

$$\approx \begin{cases} c\|\varepsilon\|_2 & \|\varepsilon\|_2 \ll 1 \\ 1 - \frac{c}{\|\varepsilon\|_2} e^{-\frac{1}{8}\|\varepsilon\|_2^2} & \|\varepsilon\|_2 \gg 1 \end{cases}$$

Proof by picture: Densities for X and Y wrt. Leb. measure look like



Application to directed last passage percolation with Gaussian weights



$T(0,n)$ = maximal sum of weights over all right-up paths from $(0,0)$ to (n,n).

Here we take the weights to be IID $N(0,1)$.

Claim: $\mathrm{Var}(T(0,n)) \geq c \log n$.

Proof: Let (η_e) be the edge weights. Define new weights $\tilde{\eta}_e = \eta_e + \delta_e$, where (δ_e) are deterministic and depending only on $|e| :=$ distance of e from $(0,0)$.

Define $\tilde{T}(0,n)$ to be like $T(0,n)$ but with the weights $(\tilde{\eta}_e)$.

Clearly, $\tilde{T}(0,n) \geq T(0,n) + \sum_{|e|=1}^{2n} \delta_{|e|}$

Since the old geodesic is a candidate for the optimal path (and actually it equals the optimal path)

Since the ... the optimal path (and actually it equals the optimal path).

Now consider the event

$$A = \{|T(0,n) - \mathbb{E}(T(0,n))| \le t\}.$$

Then $|\mathbb{P}(\eta \in A) - \mathbb{P}(\tilde{\eta} \in A)| \le d_{TV}(\mathcal{L}(\eta), \mathcal{L}(\tilde{\eta}))$.

If we take t to be, say, two standard deviations of $T(0,n)$ then by Chebyshev's ineq., $\mathbb{P}(\eta \in A) \ge \frac{3}{4}$.

Thus, if $d_{TV}(\mathcal{L}(\eta), \mathcal{L}(\tilde{\eta})) \le \frac{1}{2}$, say, then A is likely also for $\tilde{\eta}$, whence $\sum_{|e|=1}^{2n} \delta_{|e|} \le 2t =$ four standard deviations of $T(0,n)$.

Conclude: $std(T(0,n)) \ge c \sum_{|e|=1}^{2n} \delta_{|e|}$ whenever $d_{TV}(\mathcal{L}(\eta), \mathcal{L}(\tilde{\eta})) \le \frac{1}{2}$.

Choice of δ_e: Choose $\delta_{|e|} = \frac{c}{|e|\sqrt{\log n}}$.

Then $d_{TV}(\mathcal{L}(\eta), \mathcal{L}(\tilde{\eta})) \approx \|\delta_e\|_2 = \sum_{|e|=1}^{2n} \delta_{|e|}^2 \cdot |e| \le \frac{1}{2}$.

$$\Rightarrow std(T(0,n)) \ge c \sum_{|e|=1}^{2n} \delta_{|e|} = c\sqrt{\log n}.$$

Mermin-Wagner style arguments