LARGE DEVIATIONS at WORK

Ofer Zeitouni, University of Minnesota and Technion

Abel lecture, May 23 2007

Plan

1. Application I – Entropic repulsion and maxima of Gaussian Free Field via large deviations.

2. Application II – Multi-User Communication, capacity, and spherical integrals via large deviations for random matrices.

The Gaussian distribution (due to De-Moivre):

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2/2\sigma^2}$$

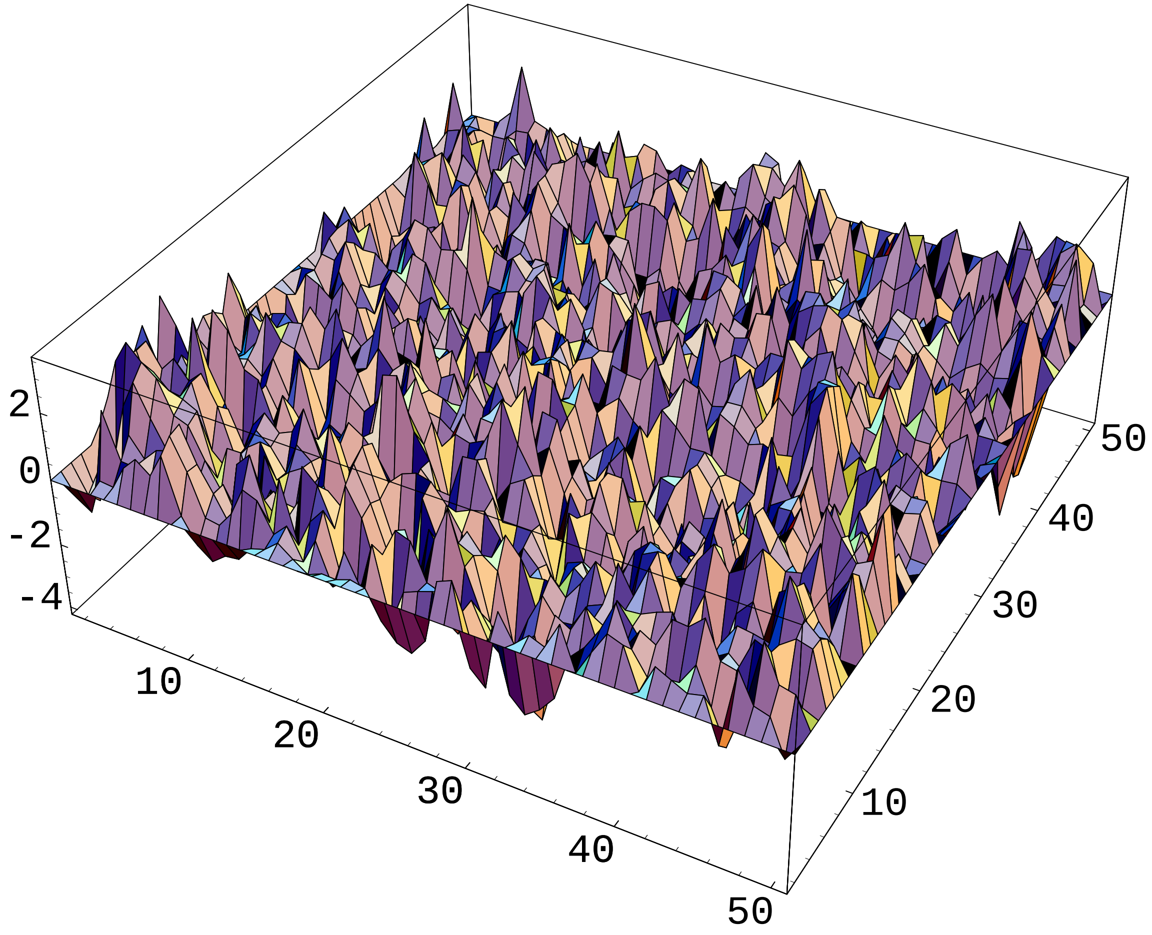
Common in nature (central limit theorem):
sums of independent variables, normalized,
tends to Gaussian

Multivariate: (X_k) has law, with covariance Λ and zero mean

$$p(x_1, \ldots, x_n) = \frac{1}{(2\pi \det \Lambda)^{n/2}} e^{-\frac{1}{2}\mathbf{x}^T \Lambda^{-1} \mathbf{x}}, \quad \mathbf{x} = (x_1, \ldots, x_n)$$

Gaussian field: k is replaced by $z \in \mathbb{Z}^d$

The Gaussian massless free field (harmonic crystal)



(courtesy of Scott Sheffield)

Models random surfaces, geophysical data, . . .

Definition of the GFF: $V = [0,1]^d$, $V_N = NV \cap \mathbb{Z}^d$

Joint density: $p(X_z, z \in V_N) = \frac{1}{Z_N} e^{-\frac{1}{2d}\sum_{i\sim j}(X_i - X_j)^2}$.

$X_i = 0$ for $i \in \partial V_N$

Gibbsian (Markovian: law of X_x depends only on neighbors)

For $d \geq 3$, extends to a Gaussian field on all of $\mathbb{Z}^d$.

Covariance equals Green function G_N of simple random walk killed at exit from V_N:

$$E(X_x, X_y) = G_N(x,y) = E_x\left(\sum_{i=0}^{\tau_{V_N}} \mathbf{1}_{S_i = y}\right) .$$

GFF is model of particles in interaction: at each vertex $z \in \mathbb{Z}^d$, put a particle, height is value of the GFF.

Rough paradigm: with many interacting particles, typical configuration is a balance between the probability of a specific configuration (as measured by the strength of the interaction) and number of configurations nearby (as measured by an entropy, or large deviations exponent).

GFF can be used to model interface, and also a droplet on a hard wall, leading to the:

Question: how does field look like if constrained to be positive? (motivation: interface model Lebowitz-Maes '87.)

Let $\Omega_N^+ = \{X_k \geq 0, k \in V_N\}$

Theorem ($d \geq 3$) There is a constant $C > 0$ such that for all $\epsilon, \delta > 0$,

$$\lim_{N\to\infty} \inf_{x\in V_N: d(x,\partial V_N)>\delta N} P\left(\left|\frac{X_x}{\sqrt{\log N}} - 4G(0,0)\right| > \epsilon \mid \Omega_N^+\right) = 1\,.$$

For $d = 2$, replace $\sqrt{\log N}$ by $\log N$. For $d = 1$, scale is $\sqrt{N}$, but no tightness

Bolthausen, Deuschel, Giacomin, Z. - '95–'01. Related to wetting phenomena and structure of droplets - Caputo, Ioffe, Velenik, Funaki, Ben Arous . . .

How does level grow? Key is large deviations estimate: for $d \geq 3$,

$$\lim \frac{1}{N^{d-2}\log N} \log P(\Omega_N^+) = -G(0,0)C$$

Constant C is a multiple of capacity of V under Newtonian potential

$$\text{cap}(V) = \sup\{2\mu(V) - \int_V \int_V \frac{1}{|x-y|^{d-2}} \mu(dx)\mu(dy) : \mu \text{ positive Radon measure on } V\}$$

More precisely, key is to understand what change of measure will create Ω_N^+. If one looks at new Gaussian field P_N with mean $\sqrt{a \log N}$ on V_N and same covariance as the free field, then relative entropy is

$$H(P_N|P) \sim \frac{a}{2}\langle \mathbf{1}_{V_N}, G_N^{-1}\mathbf{1}_{V_N}\rangle \log N$$

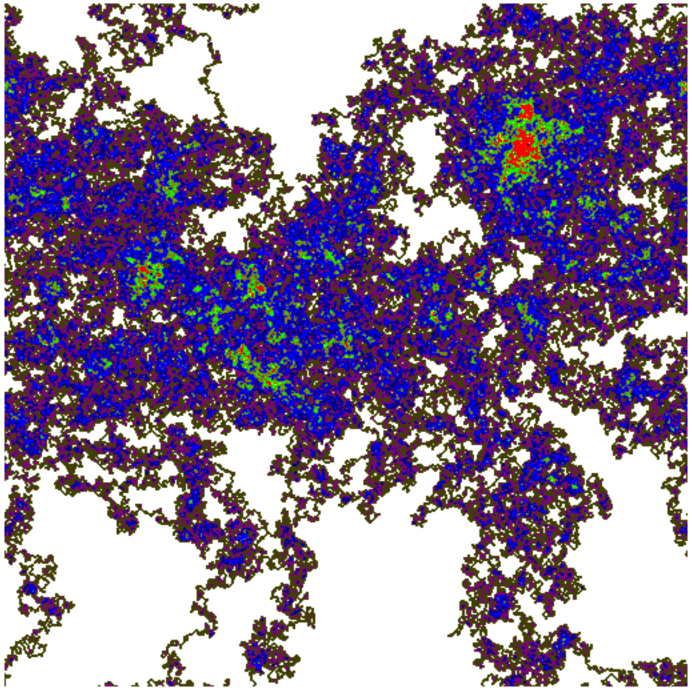
The scaling of the inner product gives N^{d-2} multiplied by the capacity. Under the new field, Ω_N^+ is typical, and one has entropic repulsion to the level announced.

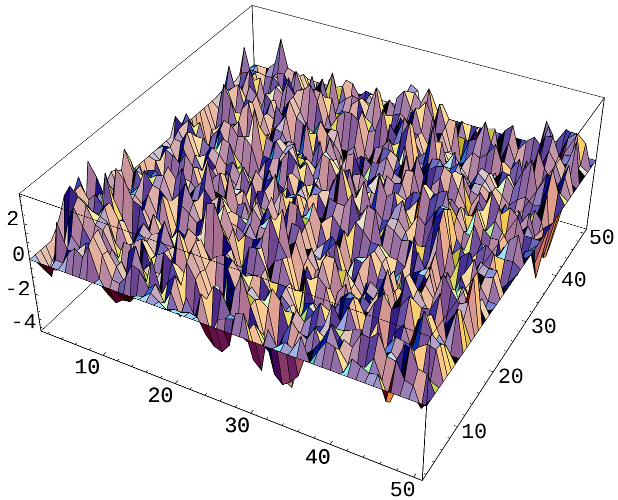
LD for empirical measure at critical scale N^{d-2} (not N^d, due to Donsker-Varadhan) - reminiscent of LD in phase transition regime (Bolthausen-Deuschel '94)

Not quite right constant, but close.

$d = 2$ is more delicate, and requires a multi-scale analysis. The extrema of the Gaussian free field possess a fractal structure, and behave like the structure of points visited often by random walk on a torus.

(Daviaud ’05; builds on Bolthausen-Deuschel-Giacomin ’01 and Dembo-Peres-Rosen-Z. ’00-’04; Precursors to the work of DPRZ can be traced back to Varadhan’s appendix to Symanzik (1969))





Figures courtesy of Raissa D’Souza and Scott Sheffield.

See Sheffield’s web site for his work with Schramm on relation of 2DGFF to SLE.

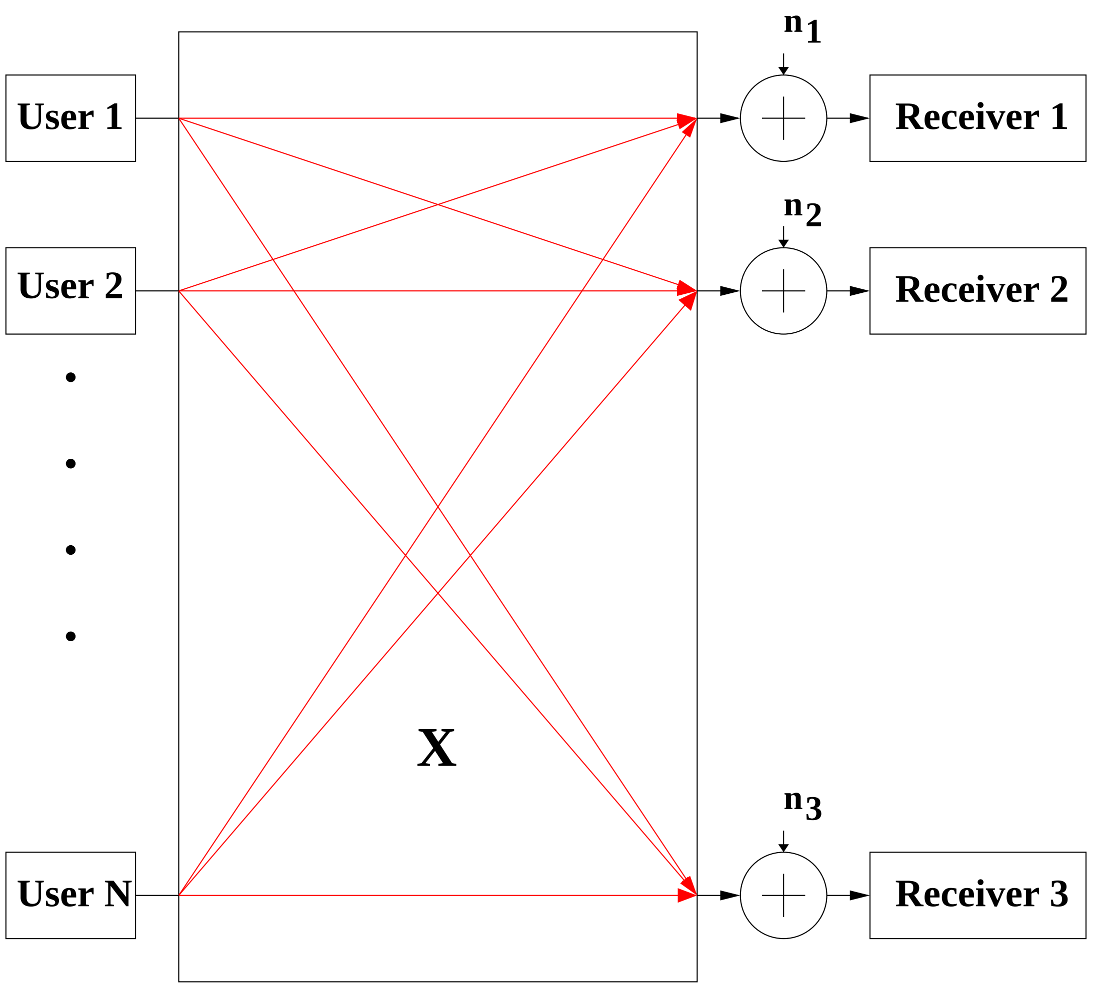
Multi-User Communication N transmitters, M receivers, channel:

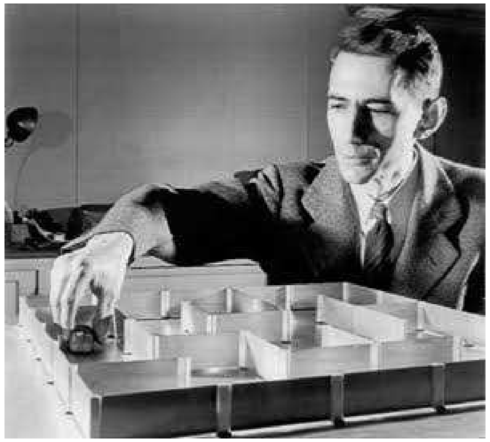
$$\mathbf{y} = X\mathbf{u} + \mathbf{n}$$

X – channel matrix, $X_{i,j}$ the (complex) gain between transmitter j and receiver i.

$\mathbf{n}$ – a vector of independent Gaussian noises.

$\mathbf{u}$ – vector of transmitted signals, subject to constraint that average power $Eu_j^2 \leq P_j$.


n1
n2
n3
User 1
User 2
User N
X
Receiver 1
Receiver 2
Receiver 3

 Under appropriate assumptions (receivers know channels, users do not), the **Shannon** capacity (i.e., maximal rate of transmition of information on the channel) is

$$C_M = \mathrm{tr}\log(I + XPX^*) = \sum \log(1 + \lambda_i)$$

where λ_i are eigenvalues of XPX^*.

C_M **is random!** But if M, N are large, maybe averaging phenomena?

Capacity: $C_M = \operatorname{tr}\log(I + XPX^*) = M\langle L_M, \log(1+x)\rangle$ where

$$L_M = \frac{1}{M}\sum_{i=1}^{M} \delta_{\lambda_i}\,.$$

Depends on the spectrum of random matrix XPX^*, a *Marcenko-Pastur* matrix.

Fact: as $N \to \infty$ with $N/M \to \alpha$, if $M^{-1}\sum \delta_{P_i} \to \mu$, then $L_M \to \sigma_{\alpha,\mu}$. (Equation for the Stieltjes transform of $\sigma_{\alpha,\mu}$).

Then, $C_M/M \to \langle \sigma_{\alpha,\mu}, \log(1+x)\rangle =: c_0$.

Question (*outage probability*): for $c < c_0$, $P(C_M/M < c) \sim$?

Digress to random matrix theory.

Random Matrices

GUE: random Hermitian matrix, i.i.d. entries - Gaussian, zero mean, $E|X_{i,j}|^2 = 1/N$.

Joint distribution of eigenvalues:

$$p(\lambda_1, \ldots, \lambda_N) = \frac{1}{Z_N} \Delta_N(\lambda)^2 \exp(-N \sum \lambda_i^2/2)$$

Here, Δ_N is Vandermonde determinant:

$$\Delta_N(\lambda) = \prod_{i<j} (\lambda_i - \lambda_j)$$

To link with statistical mechanics, if each λ_i is a "particle", then joint law gives a formula for the interaction between particles.

In terms of empirical measure $L_N = N^{-1} \sum \delta_{\lambda_i^{(N)}}$:

$$p(\lambda_1, \ldots, \lambda_N) = \frac{1}{Z_N} \exp\left[-N^2\langle L_N, x^2/2\rangle + N^2\langle L_N, \log|x-y| L_N\rangle\right]$$

$\langle \nu, f(x,y)\nu\rangle = \iint \nu(dx)\nu(dy) f(x,y)$

Large deviations What is probability $P(L_N \in A)$?

LDP:

$$N^{-2}\log P(L_N \sim \nu) \sim -I(\nu)\,,$$

where $I(\nu) = \langle \nu, x^2/2\rangle - \sum(\nu) + C$,

$$\sum(\nu) = \iint \nu(dx)\,\nu(dy)\log|x-y| \quad \texttt{Voiculescu's entropy}$$

Precisely: equip space of probability measures on $\mathbb{R}$ with weak topology. Then, with $I(A) = \inf_{\nu\in A} I(\nu)$,

$$-I(A^o) \le \liminf N^{-2}\log P(L_N \in A) \le \limsup N^{-2}\log P(L_N \in A) \le -I(\bar{A})\,.$$

(Ben Arous-Guionnet 1997)

Note surprising N^2 scale (contrast with standard LDP for sum of i.i.d.).

That correct scale is N^2 follows from concentration of measure arguments, and is not limited to Gaussian entries Guionnet-Z. ’00

I is strictly convex, hence:

Corollary: $L_N \to \sigma$ where $\sigma = \frac{1}{2\pi}\sqrt{4 - x^2}dx$ is semi-circle, solution of $I(\sigma) = 0$ (Wigner).

Wishart matrices: $Y = XX^*$, X - $M \times N$ i.i.d. Normal(1,$1/N$).

Joint distribution of eigenvalues is

$$p(\lambda_1, \ldots, \lambda_N) = \frac{1}{Z_N} \Delta_N(\lambda)^2 \prod \lambda_i^{(\alpha-1)M} \exp(-N \sum \lambda_i/2)$$

Similar LDP as in GUE case (scale N^2). Resolves capacity question when $P = I$.

Large deviations of spectrum of matrix ($N = \alpha M$)

$$\underset{\substack{\uparrow\\(N\times M)}}{X_N}\ \underset{\substack{\uparrow\\(M\times M)}}{P_N}\ X_N^*$$

iid, $\mathcal{N}\left(0, \frac{1}{N}\right)$ entries

diagonal, $\frac{1}{M}\sum_{i=1}^{M} \delta_{P_N(i,i)} \xrightarrow{w} \mu_0$

Joint distribution of eigenvalues:

$$\frac{1}{Z}\Delta(\lambda_i)^{\beta} \prod_i \lambda_i^{\frac{\beta}{2}(N-M+1)-1}\ e^{-N\sum_1^N \lambda_i} . \int e^{-\frac{N}{2}\mathrm{tr}(UP_N^{-1}U^*D_\lambda)}\ \mu_M(dU)$$

μ_M – Haar on $U_M(\beta = 2)$ or $O_M(\beta = 1)$

$\Delta(\lambda) = \prod_{i<j} |\lambda_i - \lambda_j|$ (Vandermonde)

Essential role played by integrals

$$\mathsf{I}Z = \int_{\{O_N, U_N\}} \mathsf{exp}\Big\{N\,\mathsf{tr}(A_N U_N B_N U_N^*)\Big\}\mu_N(dU_N)$$

Indeed, once spectrum of matrices in model is frozen, integration over angular (=spherical) variable leaves these:

Itzykson-Zuber-Harish Chandra Integrals

Appear in multi-matrix models in physics, enumeration of random surfaces, . . .

Itzykson-Zuber-Harish Chandra Integral

Basically, need the following:

If A_N, B_N are diagonal

$$\frac{1}{N}\sum_{i=1}^{N}\delta_{A_N(i,i)} \xrightarrow{w} \mu_A, \quad \frac{1}{N}\sum_{i=1}^{N}\delta_{B_N(i,i)} \xrightarrow{w} \mu_B$$

then

$$\frac{1}{N^2}\log\Big(\text{IZ}\,(A_N, B_N)\Big) \longrightarrow J(\mu_A, \mu_B)$$

+ info on J

Outage probability can then be written in terms of J.

Itzykson-Zuber-Harish Chandra Integrals

$$I_N(A_N, B_N) = \int_{U_N} \mathsf{exp}\Big\{ N\, \mathsf{tr}(A_N U_N B_N U_N^*) \Big\} \mu_N(dU_N)$$

A formula of Harish Chandra:

$$I_N = \frac{\mathsf{det}\{\mathsf{exp}\, N A_N(i) B_N(j)\}}{\Delta(A_N)\Delta(B_N)}, \quad \frac{1}{N^2} \log \Delta(A_N) \longrightarrow \frac{1}{2} \sum(\mu_A)$$

But $\mathtt{det}$ *has oscillating terms, and is* <u>*not*</u> *dominated by largest.*

Stochastic Calculus, Large Deviations

Suppose we knew how to compute, with

$$X_N - \text{ Wigner } \begin{pmatrix} X_N(i,i) \sim N\left(0, \frac{1}{N}\right) \\ X_N(i,j) \sim N\left(0, \frac{1}{2N}\right) \\ \text{symmetric, independent} \end{pmatrix}$$

$$D_N - \text{ diagonal, deterministic, } \frac{1}{N}\sum_{i=1}^{N} \delta_{D_N(i,i)} \xrightarrow{w} \mu_D$$

Probability of

$$\left(\frac{1}{N}\sum \delta_{\lambda_i(X_N+D_N)} \sim \mu\right)$$

Then,

$$e^{+N^2[J(\mu_X,\mu_D)-\langle\mu_X,\frac{x^2}{2}\rangle+\beta\sum(\mu_X)]}$$

$$\sim \int_{\widehat{\mu}_{X_N}\sim\mu_X} e^{-\frac{N^2}{2}\langle\widehat{\mu}_{X_N},x^2\rangle}\Delta(\lambda(X_N))^\beta \cdot I(X_N,D_N)\prod d\lambda_i(X_N)$$

$$= \iint_{\widehat{\mu}_{X_N}\sim\mu_X} e^{-\frac{N}{2}\mathrm{tr}(UD_{\lambda(X_N)}U^*-D_N)^2} d\mu_N(U)\Delta(\lambda_i)^\beta\prod d\lambda_i\cdot e^{\frac{N^2}{2}\langle x^2,\widehat{\mu}_{D_N}\rangle}$$

$$\sim e^{\frac{N^2}{2}\langle x^2,\widehat{\mu}_{D_N}\rangle}P(\widehat{\mu}_{Y_N}\sim\mu_X)\cdot Z_N$$

where $Y_N = X_N + D_N$ and $D_{\lambda(X_N)}$ is diagonal with entries = eigenvalues.

So: all boils down to computing LDP for empirical measure of

$$X_N + D_N \, .$$

This is a hard problem.

Make it harder: Compute LD of *trajectory* of empirical measures

$$X_N(t) + D_N$$

$(X_N(t))_{i,j}$: $\frac{1}{\sqrt{N}} BM(t)$, independent!

Idea: use stochastic calculus. Motivated by methods from hydrodynamic limits

Let $L_N^Y =$ empirical measure of $Y = X_N(t) + D_N$

f – test function, $f(t,x)$ smooth

Ito's formula:

$$d\langle L_N^Y, f\rangle = \langle \mathcal{L}f, L_N^Y\rangle dt + M_t + \text{ negligible terms } + \langle \mathcal{G} f L_N^Y, L_N^Y\rangle$$

M_t is martingale, bracket

$$\langle M_t\rangle = \int_0^t \langle |\partial_x f|^2, L_N^Y(s,dx)\rangle ds$$

$\Rightarrow \qquad \exp\left[M_t - \frac{1}{2}\langle M_t\rangle\right]$ are martingales.

Here:

$$\mathcal{G}f = \frac{\partial_x f(x) - \partial_x f(y)}{x - y}$$

So, if one defines

$$S^{0,t}(\nu, f) = \langle f, \nu_t \rangle - \langle f, \nu_0 \rangle - \int_0^t \langle \dot{f}, \nu_\theta \rangle d\theta - \frac{1}{2} \int_0^t \langle \mathcal{G}f(\theta, x, y) \nu_\theta, \nu_\theta \rangle$$

Then

$$M(L_N(\cdot), f) = \exp\left\{ N^2 (S^{0,t}(L_N, f) - \frac{1}{2} \int_0^t \langle f_x^2, \nu_\theta \rangle d\theta) \right\}$$

is a martingale.

A LD upper bound:

$$P\Big(L_N(\cdot) \in B(\nu_\cdot, \delta)\Big) = E\left(\mathbf{1}_{L_N(\cdot)\in B(\nu_\cdot,\delta)} M(L_N(\cdot), f)\, M^{-1}(L_N(\cdot), f)\right)$$

$$\leq e^{-N^2(S^{0,1}(\nu,f) - \frac{1}{2}\int_0^1 \langle (f')^2, \nu_s\rangle ds + O(\delta))} \cdot \underbrace{EM}_{=1!}$$

$\Rightarrow$

$$\lim_{\delta\to 0} \limsup_{N\to 0} \frac{1}{N^2} \log P\Big(L_N(\cdot) \in B(\nu_\cdot, \delta)\Big)$$

$$\leq - \sup_{f\in C_b^{1,2}} \Big(S^{0,1}(\nu, f) - \frac{1}{2} \ll f', f' \gg_\nu\Big) := I(\nu_\cdot)$$

This, + exponential tightness, + lower semi continuity of $I \Rightarrow$ upper bound.

For LB:

Note first that if $I(\nu.) < \infty$ then

$$I(\nu.) = \sup_{f \in C_b^{1,2}} \left(S^{0,1}(\nu., f) - \frac{1}{2} \ll f', f' \gg_\nu \right)$$

$$= \frac{1}{2} \sup_{f \in C_b^{1,2}} \frac{S^{0,1}(\nu, f)}{\ll f', f' \gg_\nu}$$

$\Rightarrow \quad S^{0,1}(\nu, f)$ is bounded linear functional

$\Rightarrow \quad S^{0,1}(\nu, f) = \ll f', g \gg_\nu$, some g .

Define new measure

$$E_g(\cdot) := E\Big(\mathcal{M}(g, L_N)\cdot\Big)$$

(uses Girsanov on each $X_{ij}(t)$, g is new drift, or *“magnetic field”*).

Limit points of $L_N(\cdot)$ (exist by compactness) satisfy

$$S^{0,1}(\nu, f) = \ll f', g \gg_\nu, \quad \forall f \in C_b^{1,2} \qquad (*)$$

<u>If</u> sol. to (*) is unique:

$$P\Big(L_N \in B(\nu., \delta)\Big) = E_g\Big(\mathbf{1}_{L_N \in B(\nu.,\delta)}\, M^{-1}(L_N, g)\Big)$$

$$\geq e^{-N^2(S^{0,1}(\nu,g) - \frac{1}{2}\ll g,g\gg_\nu)} \cdot E_g\Big(\mathbf{1}_{L_N \in B(\nu,\delta)}\Big) \cdot e^{-N^2 O(\delta)}$$

$$\geq e^{-N^2(I(\nu)+O(\delta))} \cdot \underbrace{E_g\Big(\mathbf{1}_{L_N \in B(\nu,\delta)}\Big)}_{\downarrow N\to\infty \atop 1}$$

$\Rightarrow$

$$\lim_{\delta\to 0}\ \liminf_{N\to\infty}\ \frac{1}{N^2}\ \log P\Big(L_N \in B(\nu,\delta)\Big) \geq -I(\nu)$$

for ν with (*) having a unique solution.

- *Cabanal-Duvillard + Guionnet (2001):*

  Above works for "smooth" $\nu_\cdot$, with g s.t. Fourier transform of g decays exponentially.

- *Guionnet + Z. (2002, 2004):*

  Extends to ν with

  $$\sup_t \langle \nu_t, x^{5+\varepsilon} \rangle < \infty$$

  Proof not trivial: involves approximations + free probability

  (+ Regularization with Cauchy).

Putting it all together

$\Rightarrow$ IZ integral given (up to known terms) by solution to variational problem

$$\inf\Big\{I(\nu.) : \quad \nu_0 = \mu_A,\ \nu_1 = \mu_B\Big\}$$

Formally: if μ_A, μ_B have smooth densities, (*) leads to $\nu(dx) = \rho_t(x)dx$ with

$$\partial_t \rho_t = -\partial_x(\rho H \rho) - \partial_x(g\rho)$$

Here

$$H\rho = PV \int \frac{1}{x-y}\rho(y)\,dy$$

Defining

$$\pi = \int \log(x-y)\rho(y,t)dt + \int_x^\infty g(y,t)\,dy$$

get

$$\begin{cases} \partial_t \pi = -\frac{1}{2}(\partial_x \pi)^2 + \frac{\pi^2}{2}\rho^2 \\ \\ \partial_t \rho = -\partial_x(\rho \partial_x \pi) \end{cases}$$

Equations suggested by Matytsin.

Congratulations, Raghu and thank you!