Lecture 2. The stochastic localization process

Informal definition. The original result of Kannan, Lovasz and Simonovits from 1995, namely $C_P(\mu) \lesssim \operatorname{tr}(\operatorname{cov}(\mu))$ for the uniform measure on a convex set, was obtained by a localization method, which, very roughly speaking, consisted in reducing the problem to a one dimensional problem by truncating the measure by affine hyperplanes repeatedly. The method of Eldan is inspired by this localization technique but as a number of key differences

- Instead of being truncated the measure is perturbed multiplicatively by a Gaussian factor, of the form $\exp\left(c + x \cdot \theta - \delta|x|^2\right)$,
- The relevant direction θ is chosen randomly, according to a certain distribution for which the above Gaussian factor is 1 on average,
- This operation is performed in continuous time rather than discrete time.

Let us now spell out the actual definition of Eldan's process. We are given a probability measure on $\mathbb{R}^n$, satisfying some mild moment conditions which we do not want to specify for now, and a standard Brownian motion (W_t) on $\mathbb{R}^n$. We consider the following infinite system of SDE whose unknown is the family (p_t) of functions from $\mathbb{R}^n$ to $\mathbb{R}_+$:

$$\begin{cases} p_0(x) = 1 \\ dp_t(x) = p_t(x)\,(x - a_t) \cdot dW_t, \end{cases}$$

where a_t is the barycenter $p_t(x)\mu(dx)$, namely

$$a_t = \frac{\int_{\mathbb{R}^n} x \cdot p_t(x)\,\mu(dx)}{\int_{\mathbb{R}^n} p_t(x)\,\mu(dx)}.$$

Note that we have only one Brownian motion (W_t) which is used for every x. Actually in Eldan's original paper [3] the process is slightly more intricate than that. Here we consider the simplified version that was introduced by Lee and Vempala [7].

Since we have an equation for each x and they are all coupled together by the condition on the barycenter, it is not at all clear that such a process should actually exists. Let us leave that matter aside for now, and go on with the main properties of the process. The barycenter condition ensures that the total mass of $p_t d\mu$ remains constant. Indeed, at least formally we have

$$d\int_{\mathbb{R}^n} p_t(x)\,\mu(dx) = \int_{\mathbb{R}^n} dp_t(x)\,\mu(dx) = \left(\int_{\mathbb{R}^n} (x - a_t) p_t(x)\,\mu(dx)\right) \cdot dW_t,$$

which is 0 by definition of a_t. Therefore $p_t d\mu$ is a random probability measure for all time, and we call that measure μ_t from now on. The second feature is that $p_t(x)$ is a martingale for all x. In particular $\mathbb{E}p_t(x) = p_0(x) = 1$ for all x. Therefore the random measure μ_t equals μ on average

$$\mathbb{E}\mu_t = \mu.$$

The third observation is that the equation

$$dp_t(x) = p_t(x)(x - a_t) \cdot dW_t$$

can be solved explicitly. Indeed applying Itô's formula to $\log p_t(x)$ we get

$$d \log p_t(x) = (x - a_t) \cdot dW_t - \frac{1}{2}|x - a_t|^2 \, dt,$$

hence

$$p_t(x) = \exp\left(\int_0^t (x - a_s) \cdot dW_s - \frac{1}{2} \int_0^t |x - a_s|^2 \, ds \right) = \exp\left(c_t + x \cdot \theta_t - \frac{t}{2}|x|^2 \right),$$

where (c_t) and (θ_t) are certain random processes not depending on x. This shows that the density p_t of μ_t with respect to μ is just a certain Gaussian factor. The linear term and the normalizing constant are random but the quadratic term is deterministic, equal to $\frac{t}{2}|x|^2$. As a result if the original measure was log-concave then the measure μ_t is t-uniformly log-concave, almost surely. Therefore the stochastic localization of Eldan allows us to write a log-concave measure as a mixture of t-uniformly log-concave measures. Moreover this mixture is constructed by solving a certain stochastic differential equation, so that its behavior over time can be somehow controlled using Itô's formula.

A proper construction of the process. We will now give a rigorous construction of the stochastic localization process. This process was first introduce by Eldan [3] (a variant of it actually), it was used in a number of subsequent works [7, 2, 5]. The construction that we give here is somewhat original, but very much inspired by Klartag-Puttermann [6].

Start with a standard n-dimensional Brownian motion (θ_t) defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a filtration $(\mathcal{F}_t)$. This is an odd name for a Brownian motion, you'll see the reason for this choice shortly. Observe that for every fixed $x \in \mathbb{R}^n$ the process

$$\exp\left(x \cdot \theta_t - \frac{t}{2}|x|^2 \right)$$

is a martingale. Using Fubini, we deduce that given a test function f,

$$N_t := \int_{\mathbb{R}^n} f(x) \cdot \exp\left(x \cdot \theta_t - \frac{t}{2}|x|^2 \right) \mu(dx).$$

also is a martingale. In particular its expectation is what we have at time 0, namely $\int_{\mathbb{R}^n} f \, d\mu$. Let μ_t be the random probability measure on $\mathbb{R}^n$ whose density with respect to $\mu(dx)$ is proportional to $\exp\left(x \cdot \theta_t - \frac{t}{2}|x|^2\right)$ and rewrite N_t as

$$N_t = \left(\int_{\mathbb{R}^n} f(x) \, d\mu_t \right) \cdot \int_{\mathbb{R}^n} \exp\left(x \cdot \theta - \frac{t}{2}|x|^2 \right) \mu(dx).$$

Note that the second factor is just the inverse normalizing factor for μ_t, as it should be. We will interpret this factor as a change in the probability space. Fix a large but finite time horizon T. The process (D_t) given by

$$D_t = \int_{\mathbb{R}^n} \exp\left(x \cdot \theta - \frac{t}{2}|x|^2 \right) \mu(dx)$$

is a positive martingale with expectation 1. Let $\mathbb{Q}$ be the probability measure on $(\Omega, \mathcal{F})$ whose density with respect to $\mathbb{P}$ is D_T. It is easy to see that a process (X_t) defined on $[0, T]$ is a $\mathbb{Q}$-martingale if and only if the process $(X_t D_t)$ is a $\mathbb{P}$-martingale. Since (N_t) was a $\mathbb{P}$-martingale we obtain the following.

Fact 1. *Under $\mathbb{Q}$, the process (M_t) given by $M_t = \int_{\mathbb{R}^n} f d\mu_t$ is a martingale.*

Getting an Itô equation for this process is a little more involved. It relies on the Girsanov change of measure formula which we spell out now.

Proposition 2 (Girsanov change of measure)**.** *If X is a $\mathbb{P}$-local martingale on $[0,T]$ then the process $\widetilde{X}$ given by*

$$\widetilde{X}_t = X_t - \int_0^t \frac{d\langle X, D\rangle_s}{D_s}$$

is a $\mathbb{Q}$-local martingale on $[0,T]$. Moreover, $\widetilde{X}$ and X have the same quadratic variation. In particular if X is a $\mathbb{P}$-Brownian motion on $[0,T]$ then $\widetilde{X}$ is a $\mathbb{Q}$-Brownian motion on $[0,T]$.

Remark. The bracket denotes the quadratic covariation of continuous semimartingales. Note that the quadratic variation under $\mathbb{P}$ is the same as the quadratic variation under $\mathbb{Q}$. Indeed, quadratic variation is defined as the limit in probability of some expression, and this is easily seen to be left unchanged by an absolutely continuous change of probability measure.

Remark. In the statement the process X is $\mathbb{R}$-valued but the result also works for vector valued martingales by applying it to each coordinate.

Proof. This is a very standard tool in stochastic calculus, we only give a very brief sketch of proof and refer to [8, section IV.38] for more details. This amounts to proving that $\widetilde{X}D$ is a $\mathbb{P}$-martingale. But, from Itô's integration by parts formula we get

$$\begin{aligned} d(\widetilde{X}D) &= (d\widetilde{X})D + \widetilde{X}(dD) + d\langle \widetilde{X}, D\rangle \\ &= (dX)D - d\langle X, D\rangle + \widetilde{X}(dD) + d\langle X, D\rangle. \end{aligned}$$

The quadratic covariation of X and D thus cancels out and we're left with martingale increments only. □

Coming back to our situation, we see that the change of measure is of the form

$$D_t = \exp(\varphi(t, \theta_t))$$

where $\varphi \colon \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$ is the function given by

$$\varphi(t, \theta) = \log\left(\int_{\mathbb{R}^n} \exp\left(\langle x, \theta\rangle - \frac{t}{2}|x|^2 \right) \mu(dx) \right). \tag{1}$$

This is not quite essential but let us assume for simplicity that $\mathrm{e}^{x\cdot\theta}$ is integrable for all $\theta \in \mathbb{R}^n$ in which case φ is smooth on $[0,\infty[\times\mathbb{R}^n$. From Itô's formula we get

$$dD_t = D_t \, \nabla\varphi(t, \theta_t) \cdot d\theta_t.$$

Then from Gisanov, we see that the process (W_t) given by

$$W_t = \theta_t - \int_0^t \nabla\varphi(t, \theta_t)\, dt$$

is a $\mathbb{Q}$-Brownian motion. We rewrite this equation as

$$d\theta_t = dW_t + \nabla\varphi(t, \theta_t)\, dt.$$

We are now in a position to prove the following.

Fact 3. *The Itô derivative of the $\mathbb{Q}$-martingale $M_t = \int_{\mathbb{R}^n} f d\mu_t$ is given by*

$$dM_t = \left(\int_{\mathbb{R}^n} f(x)(x - a_t)\, d\mu_t \right) \cdot dW_t,$$

where $a_t = \int_{\mathbb{R}^n} x\, d\mu_t$ is the barycenter of μ_t.

Proof. First of all, by differentiating under the integral sign, we obtain $\nabla\varphi(t, \theta_t) = a_t$. We see $M_t = \int f\, d\mu_t$ as a function of t and θ_t, denoted $F(t, \theta_t)$. Thus by Itô's formula and the equation for θ_t, we have

$$dM_t = \nabla F(t, \theta_t) \cdot (dW_t + \nabla\varphi(t, \theta_t)\, dt) + \frac{1}{2}\Delta F(t, \theta_t)\, dt + \partial_t F(t, \theta_t)\, dt.$$

Differentiating under the intgral sign, we see that on the one hand,

$$\nabla F(t, \theta_t) = \int_{\mathbb{R}^n} f(x)(x - a_t)\, d\mu_t,$$

and on the other hand that

$$\partial_t F = -\nabla F \cdot \nabla\varphi - \frac{1}{2}\Delta F.$$

So the absolutely continuous part in $dF(t, \theta_t)$ cancels out and the fact is proven. $\square$

Remark. Strictly speaking this only gives a construction of the process (μ_t) on a bounded time interval $[0, T]$. This will be sufficient for our needs but let us note that one could extend this construction to the whole half-line by some abstract argument à la Caratheodory. Beware though that the change of measure is only absolutely continuous when we restrict our processes to a bounded time interval. If you're not comfortable with this remark, you can safely ignore it.

As a byproduct of this construction we obtain a simple description of the law of the process (θ_t). This observation is not present in the works of Eldan, Lee-Vempala, and Chen. Its first explicit mention is in the paper of Klartag and Puttermann.

Proposition 4. *The process (θ_t) has the same law as the process $(tX + W_t)$, where (W_t) is a standard Brownian motion, and X is a random vector having law μ independent of (W_t).*

Proof. Recall that we only work on some finite time interval $[0, T]$. By the construction of the previous proposition, the law of the process (θ_t) is absolutely continuous with respect to the Wiener measure, with density $\mathrm{e}^{\varphi(T, W_T)}$. Set $\eta_t = tX + W_t$ for every t. Conditionally on the vector X, the process (η_t) is just a Brownian motion plus a constant speed deterministic drift. As a result its law is explicit, given by a very basic version of the Cameron-Martin formula: For any test function H we have

$$\mathbb{E}(H(\eta) \mid X) = \mathbb{E}\left(H(W) \cdot \mathrm{e}^{X \cdot W_T - \frac{T}{2}|X|^2} \mid X \right).$$

Taking expectation again and using Fubini we obtain

$$\mathbb{E}H(\eta) = \mathbb{E}H(W) \cdot \mathrm{e}^{\varphi(T, W_T)}$$

which indeed shows that η also has density $\mathrm{e}^{\varphi(T, W_T)}$ with respect to the Wiener measure. $\square$

One may wonder why we did not simply set $\theta_t = tX + W_t$ and proved that the corresponding measure valued process (μ_t) is a martingale using Itô's formula. This simplified approach does not work: While the process (μ_t) is then a martingale with respect to its own filtration, it is generally not a martingale with respect to the natural filtration of the Brownian motion (W_t). Let us illustrate this issue with a simple example.

Example. In dimension 1, take μ to be the standard Gaussian measure. In that case we have an explicit formula for φ namely

$$\varphi(t,\theta) = \frac{\theta^2}{2(1+t)} - \frac{1}{2}\log(1+t),$$

which gives $\nabla\varphi(t,\theta) = \frac{\theta}{1+t}$. The equation for the tilt process (θ_t) is thus

$$d\theta_t = dW_t + \frac{\theta_t}{1+t}\,dt,$$

which can be solved explicitly:

$$\theta_t = (1+t)\int_0^t \frac{dW_s}{1+s}.$$

Observe that the barycenter (a_t) of μ_t satisfies

$$a_t = \nabla\varphi(t,\theta_t) = \frac{\theta_t}{1+t} = \int_0^t \frac{dW_s}{1+s}$$

which is indeed a martingale with respect to the natural filtration of (W_t). On the other hand if we set

$$\eta_t = W_t + tX$$

where X is a standard Gaussian variable independent of (W_t), and if we let b_t be the barycenter of the corresponding tilt of μ, then we have

$$b_t = \nabla\varphi(t,\eta_t) = \frac{\eta_t}{1+t} = \int_0^t \frac{dW_s}{1+s} + \frac{X - W_s}{(1+s)^2}\,ds.$$

The absolutely continuous part is clearly not identically 0. Therefore (b_t) is not a martingale with respect to the Brownian motion (W_t). Nevertheless, according to the latest proposition the processes (θ_t) and (η_t) should have the same law, hence also the processes (a_t) and (b_t). This identity can be seen directly in that case. Indeed, clearly (θ_t) and (η_t) both are centered Gaussian processes. Therefore it is enough to check that the two covariance structures coincide. It turns out that

$$\mathbb{E}\theta_s\theta_t = \mathbb{E}\eta_s\eta_t = st + s\wedge t.$$

We leave this computation as an exercise.

Time reversal. The description of the law of the tilt process (θ_t) and of the stochastic localization process becomes even simpler after a time reversal. Observe that for every test function f the quantity $\int_{\mathbb{R}^n} f\,d\mu_t$ is the convolution of $f\rho$ by a certain Gaussian factor, where ρ is the density of μ with respect to the Lebesgue measure. More precisely if we introduce the heat semi-group

$$P_t f(x) = \mathbb{E} f(x + B_t) = f * g_t$$

where $g_t(x) = (2\pi t)^{-n/2} \mathrm{e}^{-|x|^2/2t}$ is the density of the Gaussian measure with mean 0 and covariance $t \cdot \mathrm{Id}$, then

$$\int_{\mathbb{R}^n} f\,d\mu_t = \frac{P_{1/t}(f\rho)}{P_{1/t}\rho}\left(\frac{\theta_t}{t}\right).$$

Warning: from now on (P_t) will denote the heat semigroup, and not the Langevin semigroup associated to μ of lecture 1. Now set $s = 1/t$, observe that

$$\frac{\theta_t}{t} \sim \frac{tX + B_t}{t} = X + sB_{1/s},$$

and observe also that $\widetilde{B}_s := sB_{1/s}$ is again a standard Brownian motion (this is the time reversal property of the Brownian motion). Here the symbol $\sim$ means equality in law. This means that up to the time reversal $s = 1/t$ the process $(\int f\,d\mu_t)_{t\geq 0}$ has the same distribution as $(Q_s f(X + B_s))_{s\geq 0}$, where Q_s is the operator defined by

$$Q_s f = \frac{P_s(f\rho)}{P_s\rho}.$$

Moreover, using the fact that the heat semigroup is self-adjoint in $L^2(dx)$ it is easy to see that

$$Q_s f(X + B_s) = \mathbb{E}[f(X) \mid X + B_s].$$

Putting everything together we see that the stochastic localization process initiated from μ has the same law as the measure valued process obtained by looking at the conditional law of X given $X + B_s$ and then reversing time by setting $t = 1/s$.

Remark. It is clear from this description that this process was looked at in many other contexts. Apparently it is an important tool in filtering theory, and it is also very much related to what Bauerschmidt, Bodineau and Dagallier [1] mpcall the Polcinski equation, which is used in their recent series of works on log-Sobolev inequalities for various particles models. In any case, the way the process was used by Eldan to prove inequalities was clearly novel.

References

[1] Bauerschmidt, R.; Bodineau, T.; Dagallier, B. *Stochastic dynamics and the Polchinski equation: an introduction.* Preprint, 2023.

[2] Chen, Y., *An Almost Constant Lower Bound of the Isoperimetric Coefficient in the KLS Conjecture.* Geom. Funct. Anal. (GAFA), Vol. 31, (2021), 34—61.

[3] Eldan, R., *Thin shell implies spectral gap via a stochastic localization scheme.* Geom. Funct. Anal. (GAFA), Vol. 23, (2013), 532–569.

[4] Kannan, R., Lovász, L., Simonovits M., *Isoperimetric problems for convex bodies and a localization lemma*. Discr. Comput. Geom., Vol. 13, (1995), 541–559.

[5] Klartag, B.; Lehec, J., *Bourgain's slicing problem and KLS isoperimetry up to polylog.* Geom. Funct. Anal. 32, No. 5, 1134-1159 (2022).

[6] Klartag, B., Putterman, E., *Spectral monotonicity under Gaussian convolution.* To appear in Ann. Fac. Sci. Toulouse Math.

[7] Lee Y.T.; Vempala S.S. *Eldan's Stochastic Localization and the KLS Conjecture: Isoperimetry, Concentration and Mixing*. Preprint, 2017.

[8] Rogers, L.C.G.; Williams, D., *Diffusions, Markov processes, and martingales.* Vol. 2: Itô calculus. 2nd ed. Cambridge University Press, Cambridge (2000).