McKean–Vlasov equations with irregular drift

Alexander Veretennikov
University of Leeds & HSE & IITP Moscow

2019 Winter Colloquium of LSA NRU HSE, Dec 2-6, Snegiri, Moscow Region

This is a joint work with Yulia S. Mishura (Kyiv)

December 14, 2019

Setting

We consider solutions of the (MV–SDE) equations in R^d

$$dX_t = B[t, X_t, \mu_t]dt + \Sigma[t, X_t, \mu_t]dW_t, \qquad X_0 = x. \quad (1)$$

In general we assume non-degeneracy of $\Sigma\Sigma^*$ and $\le$linear growth in x of both B, Σ. The *true* McKean–Vlasov's case is

$$B[t, x, \mu] = \int b(t, x, y)\mu(dy), \ \Sigma[t, x, \mu] = \int \sigma(t, x, y)\mu(dy).$$

Here in the equation W is a standard d-dimensional Wiener process (or $d_1 \geq d$), b and σ – vector and matrix Borel functions of corresponding dimensions d and $d \times d$), μ_t is the distribution of the process X at t. The initial data x may be random, but independent of W. A bit more general case is

$$B[\ldots] = \psi(t, x, \int b(t, x, y)\mu(dy)), \ \Sigma[\ldots] = \phi(t, x, \int \sigma(t, x, y)\mu(dy)).$$

Thanks & presentation and publication “history”

Sincere thanks to Denis Talay, Mireille Bossy, Sima Mehri for very valuable questions which helped correct quite a few gaps in the calculus.

We started working on the topic with Yulia Mishura nearly ten years ago and showed various preliminary versions of our results at several conferences; the first presentation was in September 2010 at the “Hommage V.S. Koroliuk, I.N. Kovalenko and A.V Skorokhod” conference. Now we devote it to the memory of A.V. Skorokhod (1930 – 2011). In March 2016 we made our first preprint. Today you may find the sixth version of it at [arxiv.org/abs/1603.02212]. Our journal submission was once rejected, resubmitted, and recently declined without a notice. We are now finishing a preparation of a new submission to another journal.

Setting

Reminder: we consider a MV–SDE

$$dX_t = B(t, X_t, \mu_t)dt + \Sigma(t, X_t, \mu_t)dW_t, \qquad X_0 = x.$$

An equivalent formulation in terms of measures (densities):

$$\partial_t p(t, x, x') - L^*(t, x, \mu_t) p(t, x, x') = 0, \qquad p(0, x, x') = \delta_{x-x'},$$

with $\mu_t(A) = \int_A p(t, x, x')\, dx'$ and

$$L(t, x, \mu) = \frac{1}{2}(\Sigma\Sigma^*)(t, x, \mu)\frac{\partial^2}{\partial x \partial x} + B(t, x, \mu)\frac{\partial}{\partial x}.$$

Dobrushin's paper relates to the "non-random case" $\Sigma \equiv 0$.

Some physical references

A.A. Vlasov (1938), “On Vibration Properties of Electron Gas” (Russian). J. Exp. Theor. Phys. 8 (3): 291. http://ufn.ru/ru/articles/1967/11/f/ (also reprinted in Sov. Phys. Usp. 10 721–733 (1968)).

A. A. Vlasov (1945). Theory of Vibrational Properties of an Electron Gas and Its Applications, Sci. Rep. Moscow State Univ., v.75, Physics, book 2, Pt.1, 196p. (Russian)

A. A. Vlasov, Many-Particle Theory and Its Application to Plasma, Gordon and Breach, 1961.

Some mathematical references

- Dobrushin, R.L. *Vlasov equations*. Funct. Anal. Appl. **13**, 115–123 (1979).
- Sznitman, A.-S. *Topics in propagation of chaos*, Calcul des probabilités, Ec. d'été Saint-Flour/Fr. 1989, Lect. Notes Math. **1464**, 165–251 (1991).
- Funaki, T. *A certain class of diffusion processes associated with nonlinear parabolic equations*, Z. Wahrscheinlichkeitstheorie verw. Gebiete **67**, 331–348 (1984).
- Veretennikov, A.Yu. *On ergodic measures for McKean-Vlasov stochastic equations*, MCQMC Methods, 2004, Berlin etc. Springer, 471–486 (2006).

Mathematical refs, ctd, incomplete

T. S. Chiang, McKean-Vlasov equations with discontinuous coefficients. *Soochow J. Math.*, 20(4):507–526, 1994.

M. Bossy, D. Talay, A stochastic particle method for the McKean-Vlasov and the Burgers equation. *Math. Comp.* 66: 157-192, 1997.

S. Benachour, B. Roynette, D. Talay, P. Vallois, Nonlinear self-stabilizing processes. I: Existence, invariant probability, propagation of chaos. *Stochastic Processes Appl.* 75(2): 173-201, 1998.

A. Ganz, Approximation of equilibrium distributions of some stochastic systems with McKean-Vlasov interactions, PhD Thesis, Sophia-Antipolis, 2008.

Mathematical refs, ctd, still very incomplete

- S. Méléard
- B. Jourdain
- L. Szpruch, D. Siska, et al.
- J.-F. Jabir
- S. Menozzi, et al.
- M. Röckner et al.
- V. Kolokoltsov
- R. Carmona and F. Delarue
- S. Mehri and W. Stannat

About Dobrushin's paper

Dobrushin treats, in particular, convergence as $N \to \infty$ of the system

$$dX_t^i = A(X_t^i)dt + \frac{1}{N}\sum_{j=1}^{N} V(X_t^i - X_t^j), \quad X_0^i = x^i.$$

Assuming that "sample measures" $\mu_t^N = N^{-1}\sum^N \delta(X_t^i)$ weakly converges, $\mu_t^N \Longrightarrow \mu_t$ and denoting

$$B_\mu(x) = \int V(x - x')\mu(dx'),$$

it is shown that the limiting equation for X^i is

$$d\bar{X}_t = (A(\bar{X}_t) + B(\bar{X}_t))dt, \quad \bar{X}_0 = x^i.$$

Dobrushin, Sznitman

So, it was suggested to add a Wiener process as a white noise,

$$dX_t = B(t, X_t, \mu_t)dt + dW_t, \qquad X_0 = x.$$

A recommended review is [A.-S.Sznitman (1991)], where Dobrushin's results and technique were used rather extensively, in particular, Kantorovich–Rubinstein metrics (aka Vaserstein's metric http://www.personal.psu.edu/lxv1/). Drift b was Lipschitz in x, original McKean–Vlasov case. A version of the K-R 1-metric on $[0, T]$ may be defined as

$$\rho(\mu, \nu) = \inf_{\pi:\ \pi_1=\mu, \pi_2=\nu} \int |x^1 - x^2|^1 \wedge 1\, \pi(dx^1, dx^2)$$

Convergence in this metric is equivalent to weak convergemnce.

$\exists!$ due to Dobrushin–Sznitman

Let $\Phi: X \mapsto x + W_\cdot + \int^\cdot B(s, X_s, \mu_s) ds$. Show contraction. Let (X^1, μ^1) and (X^2, μ^2) be two solutions of the equation with unit diffusion. For simplicity assume b bounded. We have,

$$\rho_t(\mu^1, \mu^2) \le \mathbb{E} \sup_{s \le t} |X_s^1 - X_s^2|$$

$$\le \int_0^t ds\, \mathbb{E} \left| \int (b(s, X_s^1, y_s)\mu^1(dy_s) - b(s, X_s^2, y_s)\mu^2(dy_s)) \right|$$

$$\le C \int_0^t ds \inf_{\pi:\ \pi_1 = \mu_s^1, \pi_2 = \mu_s^2} \int |x^1 - x^2| \wedge 1\, \pi(dx^1, dx^2) |$$

$$= C \int_0^t \rho_s(\mu_s^1, \mu_s^2)\, ds \qquad \textit{(and further apply Gronwall).}$$

Some motivation for further work

NO general enough solution, both strong and weak!

Weak solutions: Funaki (1984) established weak existence – analogous to Skorokhod's weak existence in ordinary SDEs – yet, so far, there was no full analogue of Krylov's weak existence, except for [V. 2006] with drift measurable in x with a special case of a constant non-degenerate diffusion matrix. We establish such analogue, see the Theorem 1 below. In the "true McKean–Vlasov case" no continuity with respect to either variable is required.

Strong solution existence (NOT equivalent to pathwise uniqueness (no full Yamada–Watanabe principle)! However, see

T.Kurtz, The Yamada-Watanabe-Engelbert... EJP, 2007.

Such existence was known only under Lipschitz conditions. We establish a new theorem of this type (Theorem 1 below).

Pathwise uniqueness: in a special case of diffusion that does not depend on y and is Lipschitz with respect to the variable x only, see the Theorem 3 below. See also a recent preprint [D. Lacker 2018, arxiv][1]. Again note that no continuity with respect to y in the "true McKean–Vlasov" case is required. In the general case see [P.-E. Chaudru de Raynal 2016 arxiv] under a certain Hölder contiunity.

For weak uniqueness see [T. Funaki 1986, et al.]; for *non-uniqueness see (Tugaut, Muzychka)* of stationary measures, the latter under assumptions which severely violate linear growth and global Lipschitz. We assume linear growth in x and some Lipschitz.

[1] Who also established multi-particle convregence.

Last fresh references

- Paul-Eric Chaudru de Raynal (LAMA), Strong well-posedness of McKean-Vlasov stochastic differential equation with Hölder drift, https://arxiv.org/abs/1512.08096
- Daniel Lacker, On a strong form of propagation of chaos for McKean-Vlasov equations, arXiv:1805.04476

Theorem (1: weak existence)

1. Let the functions b and σ satisfy linear growth bound condition in x, i.e., there exist $C > 0$ such that

$$|b(s,x,y)| + \|\sigma(s,x,y)\| \le C(1+|x|), \ \ \forall\ s,x,y,$$

where $|\cdot|$ stands for Euclidean norms in R^d for b and σ.
2. Diffusion matrix σ is uniformly nondegenerate in the following sense:

$$\inf_{s,x,\mu} \inf_{|\lambda|=1} \lambda^*\left(\int \sigma(s,x,y)\mu(dy)\right)\left(\int \sigma(s,x,y)\mu(dy)\right)^*\lambda > 0.$$

Then under $\mathbb{E}|x_0|^4 < \infty$ (which could be, perhaps, possible to relax) the equation (1) has a weak solution.

Proof, $dX_t = B(t, X_t, \mu_t)dt + \Sigma(t, X_t, \mu_t)dW_t$
$B(t, x, \mu) = \int b(t, x, y)\mu(dy)$, $\Sigma(t, x, \mu) = \int \sigma(t, x, y)\mu(dy)$ ("true McK–V")

The proof is based on the Skorokhod unique probability space convergence, representation $B(s, x, \mu_t) = \hat{\mathbb{E}} b(s, x, \hat{X}_t)$, and on Krylov's estimates,

$$\mathbb{E}\int_0^T f(t, X_t)dt \le N_{(d,T,\dots)}\|f\|_{L_{d+1}},$$

for any non-degenerate Ito process (not necessarily a solution of an SDE) (with bounded coefts, then localize). This estimate is applied to a *couple* of processes,

$$\mathbb{E}\int_0^T f(t, X_t, \hat{X}_t)dt \le N_{(2d,T,\dots)}\|f\|_{L_{2d+1}},$$

where $\hat{X}_t$ is an independent copy of X_t and, hence, has the same distribution.

Lemma (1: weak existence, "case σ symmetric, with ϕ & ψ")

1. Let the functions b and σ be continuous in y and let there exist $C > 0$ such that

$$|b(s,x,y)| + \|\sigma(s,x,y)\| \le C(1+|x|), \ \ \forall\, s,x,y.$$

2. Diffusion matrix σ is symmetric, and $\phi(\tilde\sigma)$ is symmetric and uniformly nondegenerate for any symmetric and uniformly nondegenerate $\tilde\sigma$ in the following strict sense:

$$\inf_{s,x,y}\ \inf_{|\lambda|=1} \lambda^*\phi(\tilde\sigma)\lambda > 0.$$

3. $\|\phi(s,x,A) - \phi(s,x,A')\| + |\psi(s,x,B) - \psi(s,x,B')| \le C(x)(\|A - A'\| + |B - B'|)$, $C(x) = C(1+|x|^2)$. *Then under* $\mathbb{E}|x_0|^4 < \infty$ *the equation (1) has a weak solution.*

Step I, Idea of proof of the Lemma 1

B and Σ bounded, Special case $\phi(A) \equiv A$, $\psi(b) \equiv b$

1. Establish a priori moment estimates. Smooth both coefficients. Use known results for smooth coefficients from [Funaki 86]. If time allows, show how to establish existence of (weak) solutions for smooth B, Σ "by hands" via successive approximations.

2. By Skorokhod's Theorem "on single probability space and a.s. convergence", we may assume that not only $\mu^n \Longrightarrow \mu$, but also a.s. on some probability space,

$$(\tilde X^n, \tilde W^n) \to (\tilde X, \tilde W),$$

where $(\tilde X^n, \tilde W^n) \sim (X^n, W^n)$ (standard).

3. Now the task is to pass to the limit in the integral equality,

$$X_t^n = x + \int_0^t B^n(s, X_s^n, \mu_s^n)ds + \int_0^t \Sigma^n(s, X_s^n, \mu_s^n)dW_s^n.$$

Step I, Idea of proof of the Lemma 1 case "no (ϕ, ψ)"

Rem: $(B, \Sigma)^*(s, x, \mu) = \int (b, \sigma)^*(s, x, y)\mu(dy)$, bounded, & σ "elliptic"

We have, with $\hat{\mathbb{E}}$ standing for the expectation wrt $\hat{X}^n$,

$$X_t^n = x + \int_0^t \hat{\mathbb{E}} b^n(s, X_s^n, \hat{X}_s^n) ds + \int_0^t \hat{\mathbb{E}} \sigma^n(s, X_s^n, \hat{X}_s^n) dW_s^n,$$

where $(\hat{X}^n)$ is an independent copy of X^n. Hence, by Skorokhod's method we may also assume that

$$(\hat{X}^n, \hat{W}^n) \to (\hat{X}, \hat{W}), \quad \text{a.s.}$$

Now, as an example, by Krylov's inequalities, we establish

$$|\mathbb{E} \int_0^t \hat{\mathbb{E}} b^n(s, X_s^n, \hat{X}_s^n) - \hat{\mathbb{E}} b(s, X_s, \hat{X}_s)) ds|$$

$$\leq \mathbb{E} \int_0^t \hat{\mathbb{E}} |b^n(s, X_s^n, \hat{X}_s^n) - b(s, X_s, \hat{X}_s))| ds \to 0.$$

Step I, Idea of proof

How to use Krylov's inequalities

Indeed, let us find n_0 (or, rather, b^{n_0}) such that

$$\|b^{n_0}(\cdot) - b(\cdot)\|_{L_{2d+1}} < \varepsilon.$$

Then (the 3rd term is tackled via the 1st one),

$$\mathbb{E}\hat{\mathbb{E}} \int_0^t |b^n(s, X_s^n, \hat{X}_s^n) - b(s, X_s, \hat{X}_s))| ds$$

$$\leq \mathbb{E}\hat{\mathbb{E}} \int_0^t |b^n(s, X_s^n, \hat{X}_s^n) - b^{n_0}(s, X_s^n, \hat{X}_s^n))| ds \qquad | \quad \overset{Krylov}{<} \quad N\varepsilon$$

$$+\mathbb{E}\hat{\mathbb{E}} \int_0^t |b^{n_0}(s, X_s^n, \hat{X}_s^n) - b^{n_0}(s, X_s, \hat{X}_s))| ds \qquad | \quad \underset{n\to\infty}{\to} \quad 0$$

$$+\mathbb{E}\hat{\mathbb{E}} \int_0^t |b^{n_0}(s, X_s, \hat{X}_s) - b(s, X_s, \hat{X}_s))| ds. \qquad | \quad \overset{Krylov}{<} \quad N\varepsilon$$

The stochastic terms are tackled similarly. $\square$

Step II, Idea of proof of the Lemma 1 case “with ψ”

$B(s,x,\mu) = \psi(\int b(s,x,y)\mu(dy))$, bounded, & σ “elliptic”

For the “ψ case” we have,

$$X_t^n = x + \int_0^t \psi(s, X_s^n, \hat{\mathbb{E}} b^n(s, X_s^n, \hat{X}_s^n))ds + \int_0^t \hat{\mathbb{E}}\sigma^n(s, X_s^n, \hat{X}_s^n)dW_s^n$$

where $(\hat{X}^n)$ is an independent copy of X^n. Again, we may assume

$$(\hat{X}^n, \hat{W}^n) \to (\hat{X}, \hat{W}), \quad \text{a.s.}$$

Now, assuming for simplicity $|\psi(s,x,b) - \psi(s,x,b')| \le C|b - b'|$, we establish

$$|\mathbb{E}\int_0^t \psi(s, X_s^n, \hat{\mathbb{E}} b^n(s, X_s^n, \hat{X}_s^n)) - \psi(\hat{\mathbb{E}} b(s, X_s, \hat{X}_s))ds|$$

$$\le C\mathbb{E}\int_0^t \hat{\mathbb{E}}|b^n(s, X_s^n, \hat{X}_s^n) - b(s, X_s, \hat{X}_s))|ds \to 0,$$

via Krylov’s estimates similarly to the case “without ψ”.

Step II, Idea of proof of the Lemma 1 case "with ϕ"

$\Sigma(s,x,\mu) = \phi(\int \sigma(s,x,y)\mu(dy))$, bounded, & σ and $\phi(\sigma)$"elliptic"

We have, with $\hat{\mathbb{E}}$ standing for the expectation wrt $\hat{X}^n$,

$$X_t^n = x + \int_0^t \phi(\hat{\mathbb{E}}\sigma^n(s, X_s^n, \hat{X}_s^n))dW_s + \int_0^t \hat{\mathbb{E}} b^n(s, X_s^n, \hat{X}_s^n)ds,$$

with $(\hat{X}^n)$ an independent copy of X^n. We may assume

$$(\hat{X}^n, \hat{W}^n) \to (\hat{X}, \hat{W}), \quad \text{a.s.}$$

Now, assuming for simplicity
$\|\phi(s,x,A) - \phi(s,x,A')\| \le C\|A - A'\|$, we get

$$|\mathbb{E}\int_0^t \phi(\hat{\mathbb{E}}\sigma^n(s, X_s^n, \hat{X}_s^n))dW_s^n - \phi(\hat{\mathbb{E}}\sigma(s, X_s, \hat{X}_s))dW_s|^2 \to 0,$$

via Krylov's estimates similarly to the above. Indeed (see the next page), let us find n_0 (or, rather, σ^{n_0}) such that

$$\|\sigma^{n_0}(\cdot) - \sigma(\cdot)\|_{L_{2d+1}} < \varepsilon.$$

Then (the 3rd term is tackled via the 1st one),

$$|\mathbb{E}\int_0^t \phi(s, X^n_s, \hat{\mathbb{E}}\sigma^n(s, X^n_s, \hat{X}^n_s))dW^n_s - \phi(s, X_s, \hat{\mathbb{E}}\sigma(s, X_s, \hat{X}_s))dW_s|^2$$

$$\le C^2\mathbb{E}\hat{\mathbb{E}}\int_0^t |\sigma^n(s, X^n_s, \hat{X}^n_s) - \sigma^{n_0}(s, X^n_s, \hat{X}^n_s))|^2 ds \quad | \quad \overset{Krylov}{<} \quad N\varepsilon$$

$$+|\mathbb{E}\hat{\mathbb{E}}(\int_0^t \phi(s, X^n_s, \sigma^{n_0}(s, X^n_s, \hat{X}^n_s)dW^n_s$$

$$- \int_0^t \phi(s, X_s, \sigma^{n_0}(s, X_s, \hat{X}_s))dW_s)|^2 \quad | \quad \underset{n\to\infty}{\overset{Skorokhod}{\to}} \quad 0$$

$$+C^2\mathbb{E}\hat{\mathbb{E}}\int_0^t |\sigma^{n_0}(s, X_s, \hat{X}_s) - \sigma(s, X_s, \hat{X}_s))|^2 ds. \quad | \quad \overset{Krylov}{<} \quad N\varepsilon$$

Proof of Theorem, Step 3: General bdd case

Σ non-symmetric; Eqn (1): $dX_t = B[t, X_t, \mu_t]dt + \Sigma[t, X_t, \mu_t]dW_t, \ X_0 = x.$

Non-symmetric diffusion matrix may be impossible to smooth without losing non-degeneracy: e.g., if $d = 1$ and σ may change sign. Hence, there is a special trick to tackle this case (from [Krylov, Veretennikov (1976)], where it was proposed, of course, for the case of Ito equations without μ). Let

$$A[s, x, \mu] = \Sigma\Sigma^*[s, x, \mu], \quad \tilde{\Sigma}[s, x, \mu] = \sqrt{A[s, x, \mu]}$$

(a symmetric matrix square root), and assume there is a (weak or strong) solution of the SDE

$$d\tilde{X}_t = B[t, \tilde{X}_t, \mu_t]dt + \tilde{\Sigma}[t, \tilde{X}_t, \mu_t]d\tilde{W}_t, \qquad \tilde{X}_0 = x, \qquad (2)$$

with a d-dimensional WP $\tilde{W}$.

Proof of Theorem, Step 3: General bdd case

Σ non-symmetric; Eqn (1): $dX_t = B[t, X_t, \mu_t]dt + \Sigma[t, X_t, \mu_t]dW_t, \; X_0 = x.$

WLOG we may and will assume that there exists another *independent* d_1-dimensional Wiener process $(\bar W_t, \; t \geq 0)$. Let I denote a $d_1 \times d_1$-dimensional unit matrix and let

$$p[s, x, \mu] := \tilde\Sigma[s, x, \mu]^{-1}\, \Sigma[s, x, \mu]. \tag{3}$$

The matrix $\tilde\Sigma[s, x, \mu] = \sqrt{A[s, x, \mu]}$ is symmetric and

$$p^* p[s, x, \mu] = \Sigma^*[s, x, \mu](A)^{-1}[s, x, \mu]\Sigma[s, x, \mu];$$

$$p^*[s, x, \mu]p[s, x, \mu]p^*[s, x, \mu]p[s, x, \mu]$$

$$= \Sigma^*(A)^{-1}(A)A)^{-1}\Sigma[s, x, \mu] = \Sigma^*(A)^{-1}\Sigma[s, x, \mu].$$

Proof of Theorem, Step 3: General bdd case

Σ non-symmetric; Eqn (1): $dX_t = B[t, X_t, \mu_t]dt + \Sigma[t, X_t, \mu_t]dW_t, \; X_0 = x.$

Let $\bar W_t$ be a new WP of dimension d_1, and denote

$$W^0_t := \int_0^t p^*[s, \tilde X_s, \mu_s]\, d\tilde W_s$$

$$+ \int_0^t (I - p^*[s, \tilde X_s, \mu_s] p[s, \tilde X_s, \mu_s])\, d\bar W_s,$$

which is again a WP (dim d_1), and notice that

$$\Sigma[s, x, \mu] p^*[s, x, \mu] = A[s, x, \mu](A[s, x, \mu])^{-1/2} = (A[s, x, \mu])^{1/2},$$

$$\Sigma[s, x, \mu] p^*[s, x, \mu] p[s, x, \mu] = (a[s, x, \mu])^{1/2} p[s, x, \mu]$$

$$= (a[s, x, \mu])^{1/2} (A[s, x, \mu])^{-1/2} \Sigma[s, x, \mu] = \Sigma[s, x, \mu].$$

Proof of Theorem, Step 3: General bdd case

Eqn (1): $dX_t = B[t, X_t, \mu_t]dt + \Sigma[t, X_t, \mu_t]dW_t, \qquad X_0 = x.$

Now, due to the stochastic integration rules,

$$\int_0^t \Sigma[s, \tilde X_s, \mu_s]\, dW_s^0$$

$$= \int \Sigma p^*[s, \tilde X_s, \mu_s]\, d\tilde W + \int \Sigma(I - p^* p)[s, \tilde X_s, \mu_s]\, d\bar W_s$$

$$= \int (A)^{1/2}[s, \tilde X_s, \mu_s]\, d\tilde W_s = \int \tilde\Sigma[s, \tilde X_s, \mu_s]\, d\tilde W_s$$

$$= \tilde X_t - x - \int_0^t B[s, \tilde X_s, \mu_s]\, ds.$$

So, $(\tilde X, W^0)$ is a (weak) solution of the equation (1), because, since we did not change measures, μ_s is still the distribution of $\tilde X_s$ by the assumption.

Proof of Theorem, Step 4: General bdd case

Σ may be non-symmetric

The last step: how to get a solution of the equation

$$d\tilde X_t = B[t,\tilde X_t,\mu_t]dt + \tilde\Sigma[t,\tilde X_t,\mu_t]d\tilde W_t, \qquad \tilde X_0 = x,$$

with a d-dimensional WP $\tilde W$ and with

$$A[s,x,\mu] = \Sigma\Sigma^*[s,x,\mu], \quad \tilde\Sigma[s,x,\mu] = \sqrt{A[s,x,\mu]}.$$

The answer is that we take a square root $\tilde\Sigma[s,x,\mu]$ by the Cauchy formula

$$\sqrt{A[t,x,\mu]} = \frac{1}{2\pi i}\oint_\Gamma \lambda^{1/2}(\lambda - A[t,x,\mu])^{-1}\,d\lambda, \qquad (4)$$

where Γ is a contour which internal domain inclues all eigenvalues of the matrix A for each t,x,μ. This is possible if A is bounded.

Proof of Theorem, Step 4: General bdd case
Σ may be non-symmetric

The mollifier for $\sqrt{A[t,x,\mu]}$ can be constructed as follows:

$$A^n[t,x,\mu] := \frac{1}{2\pi i}\oint\limits_{\Gamma} \lambda^{1/2}\left(\lambda - \left(\hat{\mathbb{E}}\sigma(t,x,\xi^n)\right)\times\left(\hat{\mathbb{E}}\sigma(t,x,\xi^n)\right)^*\right)^{-1} d\lambda,$$

where ξ is a r.v. with a distribution μ, and ξ^n is a smoothed version of ξ, that is, a r.v. with the distribution. This smoothing is only wrt the last variable, which leaves the matrix uniformly nondegenerate. Thus, the function ϕ in $\phi(t,x,\Sigma[t,x,\mu]) = \phi(t,x,\hat{\mathbb{E}}\sigma(t,x,\xi))$ here equals

$$\phi(t,x,\hat{\mathbb{E}}\sigma(t,x,\xi)) = \frac{1}{2\pi i}\oint\limits_{\Gamma} \lambda^{1/2}\left(\lambda - \left(\hat{\mathbb{E}}\sigma(t,x,\xi)\right)\times\left(\hat{\mathbb{E}}\sigma(t,x,\xi)\right)^*\right)^{-1} d\lambda.$$

Note that the function $\phi(\cdot)$ given by

$$\phi(t,x,\hat{\mathbb{E}}\sigma[t,x,\xi]) = \frac{1}{2\pi i}\oint_{\Gamma} \lambda^{1/2}\left(\lambda - \left(\hat{\mathbb{E}}\sigma(t,x,\xi)\right) \times \left(\hat{\mathbb{E}}\sigma(t,x,\xi)\right)^*\right)^{-1} d\lambda,$$

indeed, satisfies the required inequalities in the assumptions of the Lemma 1

$$\|\phi(t,x,A) - \phi(t,x,A')\| \le C(x)\|A - A'\|,$$

due to choice of Γ and by the identity for the resolvent

$$(\lambda - A)^{-1} - (\lambda - A')^{-1} = (\lambda - A)^{-1}(A' - A)(\lambda - A')^{-1}.$$

Proof of Theorem, Step 5: (The last note)

Σ may be unbounded and non-symmetric

In the unbounded case for $\sigma(\cdot)$ and $b(\cdot)$ stopping times can be used(!), and the symmetric positive definite square root $\sqrt{A[t,x,\mu]}$ can be represented by the sum

$$\sqrt{A[\ldots]} = \frac{1}{2\pi i}\sum_{k=1}^{\infty} 1(k-1 \le |x| < k) \oint_{\Gamma_k} \lambda^{1/2}(\lambda - A[t,x,\mu])^{-1} d\lambda,$$

where $\Gamma_k = \{\lambda \in \mathbb{C} : \ |\lambda| = \sup_{t,x,\mu:\, |x|\le k} \|A[t,x,\mu]\| + 1\}$, where the contour $\Gamma_k \subset \mathbb{C}$ in the complex plane is chosen in a way so that its interior contains all the eigenvalues of the (elliptic) matrix $A[s,x,\cdot]$ for $|x| \le i$, instead of a unique Cauchy (aka Riesz – Dunford) integral

$$\sqrt{A[t,x,\mu]} = \frac{1}{2\pi i}\oint_{\Gamma} \lambda^{1/2}(\lambda - A[t,x,\mu])^{-1}\, d\lambda.$$

Thanks

Thanks!