To start off we define the notation for the objects we will work on

$$dX_t = B(t, X_t, \mu_{X_t})dt + \Sigma(t, X_t, \mu_{X_t})dW_t, \qquad X_0 \sim \mu_0.$$

Given the flow of marginals μ_t we can fix the coefficients and linearize the SDE with the linearized coefficients $B^\mu(t,x)$ and $\Sigma^\mu(t,x)$. Using this, we may define the infinitesimal generator

$$\mathcal{A}_t^\mu = \frac{1}{2}\sum_{i,j=1}^N c_{ij}^\mu(t,x)\partial_{x_i x_j} + \sum_{i=1}^N B_i^\mu(t,x)\partial_{x_i}.$$

Given this operator, under reasonable assumptions we have the existance of $p(s,x;t,y)$ fundamental solution of

$$(\partial_s + \mathcal{A}_s^\mu)p^\mu(s,x;t,y) = 0,$$
$$(\partial_t - (\mathcal{A}_t^\mu)^*)p^\mu(s,x;t,y) = 0.$$

Having the transition density p^μ we may define the forward translation operator

$$U_\mu^{t,s}\phi(y) = \int p^\mu(s,x;t,y)\phi(x)dx,$$

whose definition may be easily extended to $\mathcal{P}^2(\mathbb{R}^N)$ due to the gaussian estimates on p (which are uniform over the choice of μ_t):

$$U_\mu^{t,s}u(y) = \int p^\mu(s,x;t,y)u(dx), \qquad u \in \mathcal{P}^2(\mathbb{R}^N).$$

Via this operator we may define

$$u_t^\mu(x) = U_\mu^{t,0}\mu_0,$$

the density of the solution of the linearized SDE via the marginal flow $(\mu_t)_{t\in[0,T]}$ with initial law μ_0. Via this density we are able to construct a new flow of marginals (the one of the solution of the linearized SDE via marginal flow μ_t and initial law μ_0):

$$\mathcal{L}_t^\mu(dy) = u_t^\mu(y)dy = \left(\int p^\mu(0,x;t,y)\mu_0(dx)\right)dy.$$

Here we will briefly state what [2] does to study the contraction properties on L^1 norm of u^μ. To start we need the identity[1] (28) of [2]:

$$U_\mu^{t,0} - U_\nu^{t,0} = \int_0^t \frac{d}{ds}U_\nu^{t,s}U_\mu^{s,0}ds = \int_0^t U_\nu^{t,s}((\mathcal{A}_s^\mu)^* - (\mathcal{A}_s^\nu)^*)U_\mu^{s,0}ds. \tag{1}$$

Then we need to observe that

$$\begin{aligned} ||U^{t,s}f||_{L^1} &= \int\left|\int p(s,x;t,y)f(x)dx\right|dy \\ &\overset{Gaussian\ estimates}{\leq} C\int\int |f(x+y)|\Gamma^+(y)dydx = C||f||_{L^1}. \end{aligned} \tag{2}$$

Also observe that

$$||(\mathcal{A}_s^\mu - \mathcal{A}_s^\nu)f||_{L^1} \leq C\sup_{t,x}\left(|c^\mu(t,x) - c^\nu(t,x)| + |B^\mu(t,x) - B^\nu(t,x)|\right)||f||_{W^{1,2}},$$

and by the fact that the operator $U_\mu^{s,0}$ is a bounded operator in $W^{1,2}$, check footnote[2] we get

$$||U_\mu^{s,0}f||_{W^{1,2}} \leq C\cdot s^{-1/2}||f||_{W^{1,2}}. \tag{3}$$

By (1), (2) and (3) we get

$$\begin{aligned} ||u_t^\mu - u_t^\nu||_{L^1} = ||U_\mu^{t,0}\mu_0 - U_\mu^{t,0}\mu_0||_{L^1} &\leq C\int_0^t s^{-1/2}ds||\mu_0||_{L^1}\sup_{t,x}\left(|c^\mu(t,x) - c^\nu(t,x)| + |B^\mu(t,x) - B^\nu(t,x)|\right) \\ &\leq C\sqrt{t}||\mu_0||_{L^1}\sup_{t,x}\left(|c^\mu(t,x) - c^\nu(t,x)| + |B^\mu(t,x) - B^\nu(t,x)|\right). \end{aligned} \tag{4}$$

[1]In my calculations I get the adjoint operator $(\mathcal{A}_s)^*$ but in Kolokoltsov's paper there is the backward one.

[2]I am unsure of this passage, we don't actually need the 0-derivative and due to gaussian estimates we should get something similar to $||\partial U\ f||_{L^1} \leq s^{-1/2}||f||_{L^1}$ and $||\partial^2 U\ f||_{L^1} \leq s^{-1}||f||_{L^1}$ where the second derivative is no longer integrable wrt s. I think something similar to (24.2.7) in dispense is happening.

Now that we briefly stated the ideas of [2] we can begin.

To leave as many doors open as possible we first define

$$I(f) := \int f(x)(u_t^\mu(x) - u_t^\nu(x))dx,$$

which at the moment may be seen as an indicator of closeness between the two densities, at a later moment we will take the sup for f in some bounded functional space like the bounded Holder functions or the bounded functions.

Now we apply this useful trick: if we write explicitly the definition of u^μ as the evaluation of p^μ on the distribution μ_0 in $I(f)$ we can change the order of integration to evaluate p^μ on the regular distribution $f(x)dx$, this is useful because it switches the operators in Kolokoltsov's formula (1) from being forward to being backwards while changing only marginally everything else. If we define the backward propagator operator

$$V_\mu^{s,t} g(y) := \int p^\mu(s, y; t, x) g(x) dx, \tag{5}$$

we can expand $I(f)$ this way

$$\begin{aligned}
I(f) &= \int\int f(x)(p^\mu - p^\nu)(0, y; t, x)\mu_0(dy)dx = \mu_0\left(\int f(x)(p^\mu - p^\nu)(0, \cdot; t, x)dx\right)\\
&= \mu_0\left(\int_0^t \frac{d}{ds}\left(\int\int p^\mu(0, \cdot; s, z)p^\nu(s, z; t, x)f(x)dxdz\right)ds\right)\\
&= \mu_0\left(\int_0^t \int\int \partial_{t_2} p^\mu(0, \cdot; s, z)p^\nu(s, z; t, x)f(x)dxdz + \int\int p^\mu(0, \cdot; s, z)\partial_{t_1} p^\nu(s, z; t, x)f(x)dxdzds\right)
\end{aligned}$$

by the fact that p is the fundamental solution for both the forward and backward PDEs

$$\begin{aligned}
&= \mu_0\left(\int_0^t \int\int (\mathcal{A}_s^\mu)^* p^\mu(0, \cdot; s, z)p^\nu(s, z; t, x)f(x)dxdz - \int\int p^\mu(0, \cdot; s, z)\mathcal{A}_s^\nu p^\nu(s, z; t, x)f(x)dxdzds\right)\\
&= \mu_0\left(\int_0^t \int (\mathcal{A}_s^\mu)^* p^\mu(0, \cdot; s, z)V_\nu^{s,t} f(z)dz - \int (\mathcal{A}_s^\nu)^* p^\mu(0, \cdot; s, z)V_\nu^{s,t} f(z)dzds\right)\\
&= \mu_0\left(\int_0^t \int p^\mu(0, \cdot; s, z)\left(\mathcal{A}_s^\mu - \mathcal{A}_s^\nu\right)V_\nu^{s,t} f(z)dzds\right)\\
&= \mu_0\left(\int_0^t V_\mu^{0,s}\left(\mathcal{A}_s^\mu - \mathcal{A}_s^\nu\right)V_\nu^{s,t} f ds\right) = \int_0^t \mu_0\left(V_\mu^{0,s}\left(\mathcal{A}_s^\mu - \mathcal{A}_s^\nu\right)V_\nu^{s,t} f\right)ds.
\end{aligned}$$

Using similar arguments it is possible to obtain also Kolokoltsov's formula (1):

$$I(f) = \int_0^t f\left(U_\nu^{t,s}\left((\mathcal{A}_s^\mu)^* - (\mathcal{A}_s^\nu)^*\right)U_\mu^{s,0}\mu_0\right)ds.$$

Now we can try to estimate $I(f)$:

$$I(f) \le \int_0^t \left|\mu_0\left(V_\mu^{0,s}\left(\mathcal{A}_s^\mu - \mathcal{A}_s^\nu\right)V_\nu^{s,t} f\right)\right|ds,$$

since μ_0 is a probability measure we can bound $\mu_0(g)$ with the uniform bound of g: $\mu_0(g) \le |g|_\infty$:

$$\le \int_0^t \sup_x \left(V_\mu^{0,s}\left(\mathcal{A}_s^\mu - \mathcal{A}_s^\nu\right)V_\nu^{s,t} f(x)\right)ds. \tag{6}$$

We observe that due to Holder's inequality and the fact that $p^\mu(0, x; s, y)dy$ is a probability measure for any fixed x we have uniformly in x

$$|V_\mu^{0,s} g(x)| = \left|\int p^\mu(0, x; s, y)g(y)dy\right| \le ||g||_{L^\infty}.$$

Thus continuing from (6) we have

$$\begin{aligned}
I(f) &\le \int_0^t ||\left(\mathcal{A}_s^\mu - \mathcal{A}_s^\nu\right)V_\nu^{s,t} f(x)||_{L^\infty} ds \le\\
&\le \sup_{s,x}\left(|B^\mu(s, x) - B^\nu(s, x)| + |c^\mu(s, x) - c^\nu(s, x)|\right)\int_0^t ||\bigtriangledown V_\nu^{s,t} f||_{L^\infty} + ||Hess\ V_\nu^{s,t} f||_{L^\infty} ds.
\end{aligned}$$

Theorem 1. *If f is a C_B^α function with $||f||_{C_B^\alpha} \le 1$ then*

1. $||Hess\ V_\nu^{s,t} f||_{L^\infty} \le \frac{C}{|t-s|^{1-\frac{\alpha}{2}}}$.

Proof.

$$\begin{aligned}
\left|\partial_{x_i x_j} V_\nu^{s,t} f(x)\right| &= \left|\int \partial_{x_i x_j} p^\nu(s,x;t,y) f(y) dy\right| \\
&\le \left|\int \partial_{x_i x_j} p^\nu(s,x;t,y)(f(y) - f(e^{(t-s)B}x)) dy\right| + \left|\int \partial_{x_i x_j} p^\nu(s,x;t,y) dy f(e^{(t-s)B}x)\right| \\
&\le \int |\partial_{x_i x_j} p^\nu(s,x;t,y)||e^{(t-s)B}x - y|_B^\alpha dy + \left|\partial_{x_i x_j} \underbrace{\int p^\nu(s,x;t,y) dy}_{=1}\right| |f(e^{(t-s)B}x)| \\
&\le \frac{C_{B,\alpha}}{|t-s|} \int \Gamma^+(t-s, x-y)|x - e^{-(t-s)B}y|_B^\alpha dy + 0 \le \frac{C_{B,\alpha}}{|t-s|^{1-\frac{\alpha}{2}}}.
\end{aligned}$$

by Lemma (A.5) of [3]. ☐

We will now define

$$d_{C_B^\alpha}(\mu,\nu) = \sup_{||f||_{C_B^\alpha} \le 1} \left|\int f(x)\left(\mu(dx) - \nu(dx)\right)\right|$$

the bounded anisotropic α-Holder distance.

Theorem 2. *The bounded anisotropic α-Holder distance metrizes weak convergence of measures. More precisely given $(\mu_n)_{n\in\mathbb{N}}$ and μ probability measures*

$$d_{C_B^\alpha}(\mu_n,\mu) \to 0 \Leftrightarrow \mu_n \overset{d}{\to} \mu.$$

Proof. The proof will be divided in two steps and is mostly taken from `https://sites.stat.washington.edu/jaw/COURSES/520s/522/HO.522.20/ch11c.pdf`

1) First we prove that

$$\mu_n \overset{d}{\to} \mu \Leftrightarrow \int f d\mu_n \to \int f d\mu,\ \forall f \in C_B^\alpha.$$

If $\mu_n \overset{d}{\to} \mu$ then equivalently $\int f\mu_n \to \int f\mu$ for any function $f \in bC$ which in particular means that it is true for any $f \in C_B^\alpha$. The converse is true because if $\int f\mu_n \to \int f\mu$ for any function $f \in C_B^\alpha$ then in particular it is true for any $f \in bLip$ which by Portmanteau's theorem implies weak convergence.

2) We will now prove that

$$\int f d\mu_n \to \int f d\mu\ \forall f \in C_B^\alpha \Leftrightarrow d_{C_B^\alpha}(\mu_n,\mu) \to 0.$$

The easy implication is the right-to-left one: indeed by comparison theorem

$$\lim_n \int f(x)\left(\mu_n(dx) - \mu(dx)\right) \le \lim_n \sup_{||f||_{C_B^\alpha} \le 1} \left|\int f(x)\left(\mu_n(dx) - \mu(dx)\right)\right| = \lim_n d_{C_B^\alpha}(\mu_n,\mu) \to 0.$$

The other way is more challenging, first by continuity from below of probability measures for any fixed $\epsilon > 0$ there exists K a compact set such that $\mu(K) > 1 - \epsilon$. Let $\mathcal{H} = \left\{f \in C_B^\alpha \mid ||f||_{C_B^\alpha} \le 1\right\}$, if we restrict each of these functions on K we have that $\mathcal{H}|_K$ is totally bounded with respect to the $||\cdot||_\infty$ norm by Ascoli-Arzelà's theorem, in particular $\exists k$ finite and $f_1, \cdots f_k \in \mathcal{H}|_K$ such that for any $f \in \mathcal{H}$ $\exists j$ such that $\sup_K |f - f_j| \le \epsilon$.

Now if we consider $d_B(x,y) = |x-y|_B$ and $K^\epsilon = \left\{x \in \mathbb{R}^N \mid d_B(x,K) \le \epsilon\right\}$ and f, f_j as before we have

$$\sup_{x\in K^\epsilon} |f(x) - f_j(x)| \le \sup_{x\in K^\epsilon} \left(|f(x) - f(y_x)| + |f(y_x) - f_j(y_x)| + |f_j(y_x) - f_j(x)|\right) \le \sup_{x\in K^\epsilon} \left(2\epsilon^\alpha + \epsilon\right) \le C_\alpha \epsilon^\alpha.$$

where y_x is a point in K such that $|x - y_x|_B < \epsilon$. C_α may be taken uniformly of ϵ as long as $\epsilon \le 1$.

Let $g(x) = \max\left(0, 1 - \frac{d_B(x,K)}{\epsilon}\right)$, evidently $g \in bLip \subseteq C_B^\alpha$ and $\mathbb{1}_K \le g \le \mathbb{1}_{K^\epsilon}$. Thus by taking n big enough we have by convergence against C_B^α functions that

$$\mu_n(K^\epsilon) \ge \int g(x)\mu_n(dx) > 1 - 2\epsilon.$$

Thus by taking $f \in \mathcal{H}$ and the associated f_j we have

$$\begin{aligned}\left|\int f(x)\,(\mu_n(dx)-\mu(dx))\right| &= \left|\int (f(x)-f_j(x))\,(\mu_n(dx)-\mu(dx))\right| + \left|\int f_j(x)\,(\mu_n(dx)-\mu(dx))\right| \\ &\leq \left|\int (f(x)-f_j(x))\mu_n(dx)\right| + \left|\int (f(x)-f_j(x))\mu(dx)\right| + \left|\int f_j(x)\,(\mu_n(dx)-\mu(dx))\right| \\ &\leq \left|\int_{K^\epsilon} (f(x)-f_j(x))\mu_n(dx)\right| + \left|\int_{(K^\epsilon)^c} (f(x)-f_j(x))\mu_n(dx)\right| + \left|\int_{K^\epsilon} (f(x)-f_j(x))\mu(dx)\right| + \\ &\quad + \left|\int_{(K^\epsilon)^c} (f(x)-f_j(x))\mu(dx)\right| + \left|\int f_j(x)\,(\mu_n(dx)-\mu(dx))\right| \\ &\leq C_\alpha\epsilon^\alpha + 4\epsilon + C_\alpha\epsilon^\alpha + 2\epsilon + \epsilon \leq C_\alpha\epsilon^\alpha,\end{aligned}$$

where the last term gets bounded by taking n big enough and by using convergence against C^α_B functions, this gives us the final result. □

Theorem 3. *For small values of T; if the coefficients of the SDE are C^α_B functions of y uniformly in (t,x) (the C^α_B norm is uniformly bounded in (t,x)) we have that the application $\mathcal{L} : C([0,T],\mathcal{P}(\mathbb{R}^N)) \to C([0,T],\mathcal{P}(\mathbb{R}^N))$ that $\mathcal{L}((\mu_t)_{t\in[0,T]}) = (\mathcal{L}^\mu_t)_{t\in[0,T]}$ is a contraction wrt the distance*

$$d((\mu_t)_{t\in[0,T]},(\nu_t)_{t\in[0,T]}) = \sup_{t\in[0,T]} d_{C^\alpha_B}(\mu_t,\nu_t).$$

Proof. We have

$$\begin{aligned}d_{C^\alpha_B}(\mathcal{L}^\mu_t,\mathcal{L}^\nu_t) &= \sup_{||f||_{C^\alpha_B}\leq 1} |I(f)| \leq \\ &\overset{Th.\ 1}{\leq} C\sup_{s,x}\left(|B^\mu(s,x)-B^\nu(s,x)| + |c^\mu(s,x)-c^\nu(s,x)|\right)\int_0^t \left(\frac{1}{|t-s|^{1-\frac{\alpha}{2}}} + \frac{1}{\sqrt{t-s}}\right) ds \\ &\leq C_T|t|^{\frac{\alpha}{2}}\sup_{s,x}\left(|B^\mu(s,x)-B^\nu(s,x)| + |c^\mu(s,x)-c^\nu(s,x)|\right).\end{aligned}$$

Now we observe that since the coefficients are uniformly C^α_B we have

$$|B^\mu(s,x)-B^\nu(s,x)| = \left|\int b(s,x,y)\mu_s(dy) - \int b(s,x,y)\nu_s(dy)\right| \leq Cd_{C^\alpha_B}(\mu_s,\nu_s),$$

where $C = ||b||_{C^\alpha_B}$, a priori it depends on (s,x) but since b uniformly C^α_B it can be taken uniformly in (s,x). It is also possible to prove that

$$|c^\mu(s,x)-c^\nu(s,x)| \leq C||\sigma||_\infty d_{C^\alpha_B}(\mu_s,\nu_s).$$

with these we can conclude that

$$d_{C^\alpha_B}(\mathcal{L}^\mu_t,\mathcal{L}^\nu_t) \leq C|t|^{\frac{\alpha}{2}}\sup_{s\in[0,t]} d_{C^\alpha_B}(\mu_s,\nu_s),$$

and thus

$$\sup_{t\in[0,T]} d_{C^\alpha_B}(\mathcal{L}^\mu_t,\mathcal{L}^\nu_t) \leq C|T|^{\frac{\alpha}{2}}\sup_{t\in[0,T]} d_{C^\alpha_B}(\mu_t,\nu_t).$$

which proves contraction for small valus of T. □

Remark 4. *This approach of having the* sup *in the distance over the space of functions of the same regularity of the coefficients of the SDE seems quite natural (b and f in the same bounded space). It doesn't seem impossible to use these types of techniques for even broader classes of coefficients as long as there are gaussian estimates.*

There is the property of tightness for the family of measures that are solution of an SDE with α-Holder coefficients and with initial law with finite p-moment:

Theorem 5. *Let $\mu_0 \in \mathcal{P}^p(\mathbb{R}^N)$. Let $p(s,x;t,y)$ be a fundamental solution of a forward Kolmogorov equation with α-Holderian coefficients so that Gaussian estimates exist. Then for any $\epsilon > 0$ there exists $K > 0$ such that*

$$\int_{B^c_K} u_t(x)dx < \epsilon, \qquad \int_{B^c_K} |x|^p u_t(x)dx < \epsilon.$$

Where $u_t(x) = \int p(0,y;t,x)\mu_0(dy)$.

Proof. The proof is a little variation of (3.2) in [2]; indeed the first inequality is proved there. Fix $\epsilon > 0$. Let $\tilde{\epsilon} > 0$ that we will fix later. Since μ_0 is a measure with finite p-moment there exists $K > 0$ such that

$$\mu_0(B_K^c) < \tilde{\epsilon}, \qquad \int_{B_K^c} |x|^p \mu_0(dx) < \tilde{\epsilon}.$$

Let $\tilde{K} > 0$ that we will fix later.

$$\begin{aligned}
\int_{|x|\geq K+\tilde{K}} |x|^p u_t(x)dx &\overset{Gaussian\ estimates}{\leq} C \int_{|x|\geq K+\tilde{K}} |x|^p \int \Gamma^+(x-\xi,t)\mu_0(d\xi)dx \\
&\leq C \int_{|\xi|\geq K, y\in\mathbb{R}^N} |y+\xi|^p \Gamma^+(y,t)\mu_0(d\xi)dy + C\int_{|\xi|\leq K, y\geq\tilde{K}} |y+\xi|^p\Gamma^+(y,t)\mu_0(d\xi)dy \\
&\leq C_p \mu_0(B_K^c) \int |y|^p \Gamma^+(y,t)dy + C_p \int_{B_K^c} |\xi|^p \mu_0(d\xi) \\
&\quad + C_p \mu_0(B_k) \int_{B_{\tilde{K}}^c} |y|^p \Gamma^+(y,t)dy + C_p \int |\xi|^p \mu_0(d\xi) \int_{B_{\tilde{K}}^c} \Gamma^+(y,t)dy
\end{aligned}$$

The first two terms get bounded by the preliminary inequalities and the fact that the Gaussian has finite p moment. The last two terms get bounded by a constant $\tilde{C}_{p,\tilde{K},T}$ that goes to 0 as $\tilde{K}$ goes to $+\infty$.

$$\leq C_{T,p}\tilde{\epsilon} + C_p\tilde{\epsilon} + C_p\tilde{C}_{p,\tilde{K},T} + C_{p,\mu_0}\tilde{C}_{p,\tilde{K},T}.$$

if we choose $\tilde{\epsilon}$ small enough and $\tilde{K}$ big enough the final result will be smaller than ϵ. We must also notice that all the estimates and the constant do not depend directly on $u_t(x)$ but on the Gaussian estimates and so they hold uniformly for the whole family of solutions. □

We may notice that the family of the marginals is bounded in $\mathcal{P}^p$ with the Wasserstein metric.

Theorem 6. *Let $(\mu_i)_{i\in\mathcal{I}}$ be the family of the marginals of solutions to SDEs with the same initial datum and α-Holderian coefficients. (Written like this is not very rigorous but for example given $(\mu_t)_{t\in[0,T]}$ the flow of marginals of a solution to an SDE as in the hypothesis we have that μ_t is an element of the family for every $t \in [0,T]$).*

Then the family is bounded as a subset of $\mathcal{P}^p(\mathbb{R}^N)$ equipped with the Wasserstein metric .

Proof. Let μ_1 and μ_2 be elements of the family. Fix $\epsilon > 0$. By theorem 5 we know that exists $K > 0$ such that

$$\int_{B_K^c} \mu_i(dx) < \epsilon, \qquad \int_{B_K^c} |x|^p \mu_i(dx) < \epsilon, \qquad i = 1,2.$$

In particular given a measure γ on $\mathbb{R}^{2N}$ with marginals μ_1 and μ_2 we have that exists $\tilde{K}$ (uniformly in γ) such that

$$\int_{B_{\tilde{K}}^c} \gamma(dy,dz) \leq \int\int_{B_K^c\times B_K^c} \gamma(dy,dz) \leq \int\int_{B_K^c\times\mathbb{R}^N} \gamma(dy,dz) = \int_{B_K^c} \mu_1(dy) < \epsilon.$$

This also works for the p-moment and we get

$$\int_{B_{\tilde{K}}^c} |y|^p\gamma(dy,dz) \leq \int\int_{B_K^c\times B_K^c} |y|^p\gamma(dy,dz) \leq \int\int_{B_K^c\times\mathbb{R}^N} |y|^p\gamma(dy,dz) = \int_{B_K^c} |y|^p\mu_1(dy) < \epsilon.$$

This means that we can bound the Wasserstein distance between the two in the following way:

$$\begin{aligned}
W^{(p)}(\mu_1,\mu_2)^p = \inf_\gamma \int\int (y-z)^p\gamma(dy,dz) &\leq \inf_\gamma \int_{B_{\tilde{K}}^c} (y-z)^p\gamma(dy,dz) + \int_{B_{\tilde{K}}} (y-z)^p\gamma(dy,dz) \\
&\leq C_p\left(\int_{B_{\tilde{K}}^c} |y|^p\gamma(dy,dz) + \int_{B_{\tilde{K}}^c} |z|^p\gamma(dy,dz)\right) + \int_{B_{\tilde{K}}} diam(B_{\tilde{K}})^p\gamma(dy,dz) \\
&\leq 2C_p\epsilon + diam(B_{\tilde{K}})^p.
\end{aligned}$$

□

The preceding theorem in particular proves that the p-moments are uniformly bounded, for this reason the following theorem is valid in our case.

Theorem 7. *Let $(\mu_i)_{i\in\mathcal{I}} \subset \mathcal{P}^p(\mathbb{R}^N)$ be a family of probability measures with tightness property as of Theorem 5 and such that the p-moments are uniformly bounded. Then for any sequence $(\mu_n)_{n\in\mathbb{N}}$ there exists a subsequence $(\mu_{n_m})_{m\in\mathbb{N}}$ and a measure $\mu \in \mathcal{P}^p(\mathbb{R}^N)$ such that*

$$\mu_{n_m} \overset{Wasserstein}{\to} \mu.$$

Proof. By theorem (6.9) of [5] we have that Wasserstein convergence in $\mathcal{P}^p$ is equivalent to weak convergence and convergence of the p-moment. We know that

$$\mu_i(B_K^c) < \epsilon, \qquad \int_{B_K^c} |x|^p \mu_i(dx) < \epsilon.$$

Thus if we define $P_i(dx) = |x|^p \mu_i(dx)$ we have that $(P_i)_{i\in\mathcal{I}}$ is a tight family of uniformly finite measures, in particular we may use a generalization of Prokhorov's theorem (Theorem 8.6.2 of [1]) and get that for any sequence there exists a subsequence P_{n_m} that converges weakly to a finite measure P.

$$P_{n_m}(\phi) \to P(\phi), \qquad \forall \phi \in C_0^\infty(\mathbb{R}^N). \tag{7}$$

By using the same theorem on μ_{n_m} we can find a new subsequence (that will still be witten as μ_{n_m}) such that μ_{n_m} converges weakly to μ.

$$\mu_{n_m}(\phi) \to \mu(\phi), \qquad \forall \phi \in C_0^\infty(\mathbb{R}^N). \tag{8}$$

Consider now $\phi \in C_0^\infty(\mathbb{R}^N)$.

$$\begin{aligned}&\int |x|^p \phi(x) \mu_{n_m}(dx) \overset{(8)}{\to} \int |x|^p \phi(x) \mu(dx) \\ &= P_{n_m}(\phi) \overset{(7)}{\to} P(\phi).\end{aligned}$$

By uniqueness of the limit we have for any $\phi \in C_0^\infty$ that $P(\phi) = \int |x|^p \phi(x) \mu(dx)$. This proves convergence of the p-moment and thus with weak convergence we have Wasserstein convergence. □

References

[1] Vladimir I. Bogachev - Measure Theory (2007)

[2] Vassili N. Kolokoltsov - Nonlinear Diffusions and Stable-Like Processes with Coefficients Depending on the Median or VaR (2013)

[3] G. Lucertini, A. Pagliarani, A. Pascucci - Optimal regularity for degenerate Kolmogorov equations in non-divergence form with rough-in-time coefficients (2024), https://doi.org/10.1007/s00028-023-00916-9

[4] Yaozhong Hu, Michael A. Kouritzin, Jiayu Zheng - Nonlinear McKean-Vlasov diffusions under the weak Hormander condition with quantile-dependent coefficients (2021), https://arxiv.org/abs/2101.04080

[5] Cédric Villani - Optimal Transport Old and New (2009)