

Generalized couplings and SFDEs

Michael Scheutzow

Technische Universität Berlin

International Seminar on SDEs and Related Topics,
November 26, 2021

Definition

P a Markov kernel on (E, \mathcal{E}) . \mathbb{P}_x law of chain starting at $x \in E$.

$$\mathcal{C}(\mathbb{P}_x, \mathbb{P}_y) := \{\xi \in \mathcal{M}_1(E^{\mathbb{N}_0} \times E^{\mathbb{N}_0}) : \xi_1 = \mathbb{P}_x, \xi_2 = \mathbb{P}_y\}$$

is the set of *couplings* of \mathbb{P}_x and \mathbb{P}_y .

Definition

P a Markov kernel on (E, \mathcal{E}) . \mathbb{P}_x law of chain starting at $x \in E$.

$$\mathcal{C}(\mathbb{P}_x, \mathbb{P}_y) := \{\xi \in \mathcal{M}_1(E^{\mathbb{N}_0} \times E^{\mathbb{N}_0}) : \xi_1 = \mathbb{P}_x, \xi_2 = \mathbb{P}_y\}$$

is the set of *couplings* of \mathbb{P}_x and \mathbb{P}_y .

Proposition

If P admits an ipm μ and, for each $x, y \in E$ there is some $\xi \in \mathcal{C}(\mathbb{P}_x, \mathbb{P}_y)$ s.t. $\xi((\zeta, \eta) : \zeta_i = \eta_i) \rightarrow 1$, then μ is unique and $P_n(x, \cdot) \rightarrow \mu$ in TV.

Definition

P a Markov kernel on (E, \mathcal{E}) . \mathbb{P}_x law of chain starting at $x \in E$.

$$\mathcal{C}(\mathbb{P}_x, \mathbb{P}_y) := \{\xi \in \mathcal{M}_1(E^{\mathbb{N}_0} \times E^{\mathbb{N}_0}) : \xi_1 = \mathbb{P}_x, \xi_2 = \mathbb{P}_y\}$$

is the set of *couplings* of \mathbb{P}_x and \mathbb{P}_y .

Proposition

If P admits an ipm μ and, for each $x, y \in E$ there is some $\xi \in \mathcal{C}(\mathbb{P}_x, \mathbb{P}_y)$ s.t. $\xi((\zeta, \eta) : \zeta_i = \eta_i) \rightarrow 1$, then μ is unique and $P_n(x, \cdot) \rightarrow \mu$ in TV.

Problems

- Sometimes, such couplings don't exist even if μ is unique. Then try to find *asymptotic coupling*.

Definition

P a Markov kernel on (E, \mathcal{E}) . \mathbb{P}_x law of chain starting at $x \in E$.

$$\mathcal{C}(\mathbb{P}_x, \mathbb{P}_y) := \{\xi \in \mathcal{M}_1(E^{\mathbb{N}_0} \times E^{\mathbb{N}_0}) : \xi_1 = \mathbb{P}_x, \xi_2 = \mathbb{P}_y\}$$

is the set of *couplings* of \mathbb{P}_x and \mathbb{P}_y .

Proposition

If P admits an ipm μ and, for each $x, y \in E$ there is some $\xi \in \mathcal{C}(\mathbb{P}_x, \mathbb{P}_y)$ s.t. $\xi((\zeta, \eta) : \zeta_i = \eta_i) \rightarrow 1$, then μ is unique and $P_n(x, \cdot) \rightarrow \mu$ in TV.

Problems

- Sometimes, such couplings don't exist even if μ is unique. Then try to find *asymptotic coupling*.
- Even asymptotic couplings are often hard to find. Try generalized (as.) couplings.

Generalized couplings

Definition

$$\hat{\mathcal{C}}(\mathbb{P}_x, \mathbb{P}_y) := \{\xi \in \mathcal{M}_1(E^{\mathbb{N}_0} \times E^{\mathbb{N}_0}) : \xi_1 \ll \mathbb{P}_x, \xi_2 \ll \mathbb{P}_y\}$$

is the set of *generalized couplings* of \mathbb{P}_x and \mathbb{P}_y .

Generalized couplings

Definition

$$\hat{\mathcal{C}}(\mathbb{P}_x, \mathbb{P}_y) := \{\xi \in \mathcal{M}_1(E^{\mathbb{N}_0} \times E^{\mathbb{N}_0}) : \xi_1 \ll \mathbb{P}_x, \xi_2 \ll \mathbb{P}_y\}$$

is the set of *generalized couplings* of \mathbb{P}_x and \mathbb{P}_y .

Example: SFDE

$$dX(t) = f(X_t) dt + g(X_t) dW(t),$$

Generalized couplings

Definition

$$\hat{\mathcal{C}}(\mathbb{P}_x, \mathbb{P}_y) := \{\xi \in \mathcal{M}_1(E^{\mathbb{N}_0} \times E^{\mathbb{N}_0}) : \xi_1 \ll \mathbb{P}_x, \xi_2 \ll \mathbb{P}_y\}$$

is the set of *generalized couplings* of \mathbb{P}_x and \mathbb{P}_y .

Example: SFDE

$$dX(t) = f(X_t) dt + g(X_t) dW(t), \quad X_0 = x \in \mathcal{C} = \mathcal{C}([-1, 0], \mathbb{R}^d),$$

Generalized couplings

Definition

$$\hat{\mathcal{C}}(\mathbb{P}_x, \mathbb{P}_y) := \{\xi \in \mathcal{M}_1(E^{\mathbb{N}_0} \times E^{\mathbb{N}_0}) : \xi_1 \ll \mathbb{P}_x, \xi_2 \ll \mathbb{P}_y\}$$

is the set of *generalized couplings* of \mathbb{P}_x and \mathbb{P}_y .

Example: SFDE

$$dX(t) = f(X_t) dt + g(X_t) dW(t), \quad X_0 = x \in \mathcal{C} = \mathcal{C}([-1, 0], \mathbb{R}^d),$$

- $X_t(s) := X(t + s), s \in [-1, 0]$.
- W is m -dim. BM

Generalized couplings

Definition

$$\hat{\mathcal{C}}(\mathbb{P}_x, \mathbb{P}_y) := \{\xi \in \mathcal{M}_1(E^{\mathbb{N}_0} \times E^{\mathbb{N}_0}) : \xi_1 \ll \mathbb{P}_x, \xi_2 \ll \mathbb{P}_y\}$$

is the set of *generalized couplings* of \mathbb{P}_x and \mathbb{P}_y .

Example: SFDE

$$dX(t) = f(X_t) dt + g(X_t) dW(t), \quad X_0 = x \in \mathcal{C} = \mathcal{C}([-1, 0], \mathbb{R}^d),$$

- $X_t(s) := X(t+s)$, $s \in [-1, 0]$.
- W is m -dim. BM

Consider

$$dX(t) = f(X_t) dt + g(X_t) dW(t), \quad X_0 = x,$$

$$dY(t) = f(Y_t) dt + g(Y_t) dW(t) + \lambda(X(t) - Y(t))dt, \quad Y_0 = y$$

Generalized couplings

Definition

$$\hat{\mathcal{C}}(\mathbb{P}_x, \mathbb{P}_y) := \{\xi \in \mathcal{M}_1(E^{\mathbb{N}_0} \times E^{\mathbb{N}_0}) : \xi_1 \ll \mathbb{P}_x, \xi_2 \ll \mathbb{P}_y\}$$

is the set of *generalized couplings* of \mathbb{P}_x and \mathbb{P}_y .

Example: SFDE

$$dX(t) = f(X_t) dt + g(X_t) dW(t), \quad X_0 = x \in \mathcal{C} = \mathcal{C}([-1, 0], \mathbb{R}^d),$$

- $X_t(s) := X(t+s)$, $s \in [-1, 0]$.
- W is m -dim. BM

Consider

$$dX(t) = f(X_t) dt + g(X_t) dW(t), \quad X_0 = x,$$

$$dY(t) = f(Y_t) dt + g(Y_t) dW(t) + \lambda(X(t) - Y(t))dt, \quad Y_0 = y$$

Under suitable conditions, for λ large, this defines a gen. coupling and $|X(t) - Y(t)| \rightarrow 0$ a.s. and uniqueness of an ipm follows from next theorem.



Theorem (S. 2020)

Let (E, d) be a metric space and P a Markov kernel on E . Assume that for every $x, y \in E$ there exists $\alpha_{x,y} > 0$ s.t. for every $\varepsilon > 0$ there exists some $\xi_{x,y}^\varepsilon \in \hat{\mathcal{C}}(\mathbb{P}_x, \mathbb{P}_y)$ s.t.

$$\xi_{x,y}^\varepsilon \left((\zeta, \eta) \in E^{\mathbb{N}_0} \times E^{\mathbb{N}_0} : \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{1}_{[0,\varepsilon]}(\rho(\zeta_i, \eta_i)) \geq \alpha_{x,y} \right) > 0,$$

where either $\rho = d$, or ρ is lsc and positive definite and E is Polish. Then P admits at most one ipm.

Theorem (S. 2020)

Let (E, d) be a metric space and P a Markov kernel on E . Assume that for every $x, y \in E$ there exists $\alpha_{x,y} > 0$ s.t. for every $\varepsilon > 0$ there exists some $\xi_{x,y}^\varepsilon \in \hat{\mathcal{C}}(\mathbb{P}_x, \mathbb{P}_y)$ s.t.

$$\xi_{x,y}^\varepsilon \left((\zeta, \eta) \in E^{\mathbb{N}_0} \times E^{\mathbb{N}_0} : \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{1}_{[0,\varepsilon]}(\rho(\zeta_i, \eta_i)) \geq \alpha_{x,y} \right) > 0,$$

where either $\rho = d$, or ρ is lsc and positive definite and E is Polish. Then P admits at most one ipm.

NB: This generalizes results in Hairer, Mattingly, S. (2011) and Kulik, S. (2017).

Main result

Theorem (S. 2020)

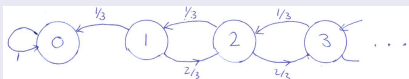
Let (E, d) be a metric space and P a Markov kernel on E . Assume that for every $x, y \in E$ there exists $\alpha_{x,y} > 0$ s.t. for every $\varepsilon > 0$ there exists some $\xi_{x,y}^\varepsilon \in \hat{\mathcal{C}}(\mathbb{P}_x, \mathbb{P}_y)$ s.t.

$$\xi_{x,y}^\varepsilon \left((\zeta, \eta) \in E^{\mathbb{N}_0} \times E^{\mathbb{N}_0} : \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{1}_{[0,\varepsilon]}(\rho(\zeta_i, \eta_i)) \geq \alpha_{x,y} \right) > 0,$$

where either $\rho = d$, or ρ is lsc and positive definite and E is Polish. Then P admits at most one ipm.

NB: This generalizes results in Hairer, Mattingly, S. (2011) and Kulik, S. (2017).

Example



Idea of proof when (E, d) is Polish

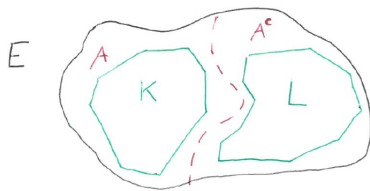
Assume there exist at least two different ipm

- Then, by ergodic decomposition theorem, there are two different ergodic ipm. These are mutually singular.

Idea of proof when (E, d) is Polish

Assume there exist at least two different ipm

- Then, by ergodic decomposition theorem, there are two different ergodic ipm. These are mutually singular.
- Let μ, ν be mutually singular ipm concentrated on A and A^c respectively. By inner regularity, for $\varepsilon > 0$ there exist compact sets $K \subset A$ and $L \subset A^c$ s.t. $\mu(K^c) < \varepsilon, \nu(L^c) < \varepsilon$. Then $d(K, L) > 0$.
- Apply Birkhoff's ergodic theorem to get a contradiction.



NB:

The assumption that (E, d) is Polish is used twice:

- Ergodic decomposition theorem
- Inner regularity of every prob. measure

NB:

The assumption that (E, d) is Polish is used twice:

- Ergodic decomposition theorem
- Inner regularity of every prob. measure

Proposition (Replacement for ergodic decomposition theorem)

If the Markov kernel P on (E, \mathcal{E}) admits at least two ipm. Then there exist two mutually singular ipm.

Replacement for ergodic decomposition theorem

Proposition (Replacement for ergodic decomposition theorem)

If the Markov kernel P on (E, \mathcal{E}) admits at least two ipm. Then there exist two mutually singular ipm.

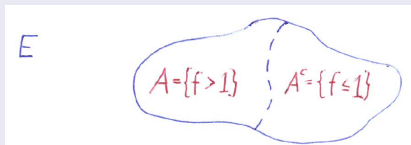
Replacement for ergodic decomposition theorem

Proposition (Replacement for ergodic decomposition theorem)

If the Markov kernel P on (E, \mathcal{E}) admits at least two ipm. Then there exist two mutually singular ipm.

Part of the proof

Assume that μ, ν are different ipm. Assume they are equivalent and $f = \frac{d\mu}{d\nu}$.



$$\int_A P(y, A^c) d\nu(y) = \int_{A^c} P(y, A) d\nu(y)$$
$$\int_A P(y, A^c) f(y) d\nu(y) = \int_{A^c} P(y, A) f(y) d\nu(y),$$

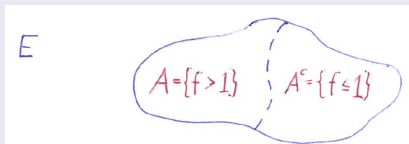
Replacement for ergodic decomposition theorem

Proposition (Replacement for ergodic decomposition theorem)

If the Markov kernel P on (E, \mathcal{E}) admits at least two ipm. Then there exist two mutually singular ipm.

Part of the proof

Assume that μ, ν are different ipm. Assume they are equivalent and $f = \frac{d\mu}{d\nu}$.



$$\int_A P(y, A^c) d\nu(y) = \int_{A^c} P(y, A) d\nu(y)$$
$$\int_A P(y, A^c) f(y) d\nu(y) = \int_{A^c} P(y, A) f(y) d\nu(y),$$

hence $P(y, A^c) = 0$ for μ -a.a. $y \in A$ and vice versa, so $\frac{1}{\mu(A)}\mu|_A$ and $\frac{1}{\mu(A^c)}\mu|_{A^c}$ are mutually singular ipm.

Proposition (Replacement for inner regularity)

Let $\mu \perp \nu$ be pm on (E, d) and $A \in \mathcal{B}(E)$ s.t. $\mu(A) = 1$, $\nu(A) = 0$. For $\varepsilon > 0$ there exist closed sets $K \subset A$ and $L \subset A^c$ s.t. $\mu(K) > 1 - \varepsilon$ and $\nu(L) > 1 - \varepsilon$ and $d(K, L) > 0$.

Proposition (Replacement for inner regularity)

Let $\mu \perp \nu$ be pm on (E, d) and $A \in \mathcal{B}(E)$ s.t. $\mu(A) = 1$, $\nu(A) = 0$. For $\varepsilon > 0$ there exist closed sets $K \subset A$ and $L \subset A^c$ s.t. $\mu(K) > 1 - \varepsilon$ and $\nu(L) > 1 - \varepsilon$ and $d(K, L) > 0$.

Idea of proof

Fix $\varepsilon > 0$. Pick *closed* sets $K \subset A$, $B \subset A^c$ s.t. $\mu(K), \nu(B) > 1 - \varepsilon$. Note:

$$B \cap (K_{1/n})^c \uparrow B.$$

Then, for n large and $L := B \cap (K_{1/n})^c$ the claim holds.

A counterexample

If P is a Markov kernel on a separable metric space (E, d) , $\rho : E \times E \rightarrow [0, \infty)$ is continuous and pos. definite such that for every pair $x, y \in E$ there exists a coupling of \mathbb{P}_x and \mathbb{P}_y s.t. $\rho(X_n, Y_n) \rightarrow 0$ a.s., then uniqueness of an ipm does *not* follow.

A counterexample

If P is a Markov kernel on a separable metric space (E, d) , $\rho : E \times E \rightarrow [0, \infty)$ is continuous and pos. definite such that for every pair $x, y \in E$ there exists a coupling of \mathbb{P}_x and \mathbb{P}_y s.t. $\rho(X_n, Y_n) \rightarrow 0$ a.s., then uniqueness of an ipm does *not* follow.

Example

Let $I \subset [0, 1]$ s.t. $\lambda^*(I) = 1$ and $\lambda_*(I) = 0$ and $J := [0, 1] \setminus I$.

$$E_1 := \{(z, 1) : z \in I\}, E_2 := \{(z, 2) : z \in J\}, E := E_1 \cup E_2.$$

$$d(x, y) = \begin{cases} |x - y|, & \text{if } (x, y) \in E_1 \times E_1 \text{ or } (x, y) \in E_2 \times E_2, \\ 1 & \text{if } (x, y) \in E_1 \times E_2 \text{ or } (x, y) \in E_2 \times E_1. \end{cases}$$

(E, d) is separable metric.

A counterexample

If P is a Markov kernel on a separable metric space (E, d) , $\rho : E \times E \rightarrow [0, \infty)$ is continuous and pos. definite such that for every pair $x, y \in E$ there exists a coupling of \mathbb{P}_x and \mathbb{P}_y s.t. $\rho(X_n, Y_n) \rightarrow 0$ a.s., then uniqueness of an ipm does *not* follow.

Example

Let $I \subset [0, 1]$ s.t. $\lambda^*(I) = 1$ and $\lambda_*(I) = 0$ and $J := [0, 1] \setminus I$.

$$E_1 := \{(z, 1) : z \in I\}, E_2 := \{(z, 2) : z \in J\}, E := E_1 \cup E_2.$$

$$d(x, y) = \begin{cases} |x - y|, & \text{if } (x, y) \in E_1 \times E_1 \text{ or } (x, y) \in E_2 \times E_2, \\ 1 & \text{if } (x, y) \in E_1 \times E_2 \text{ or } (x, y) \in E_2 \times E_1. \end{cases}$$

(E, d) is separable metric. $\rho((v, i), (w, j)) := |v - w|$ is continuous and pos. definite.

A counterexample

If P is a Markov kernel on a separable metric space (E, d) , $\rho : E \times E \rightarrow [0, \infty)$ is continuous and pos. definite such that for every pair $x, y \in E$ there exists a coupling of \mathbb{P}_x and \mathbb{P}_y s.t. $\rho(X_n, Y_n) \rightarrow 0$ a.s., then uniqueness of an ipm does *not* follow.

Example

Let $I \subset [0, 1]$ s.t. $\lambda^*(I) = 1$ and $\lambda_*(I) = 0$ and $J := [0, 1] \setminus I$.

$$E_1 := \{(z, 1) : z \in I\}, E_2 := \{(z, 2) : z \in J\}, E := E_1 \cup E_2.$$

$$d(x, y) = \begin{cases} |x - y|, & \text{if } (x, y) \in E_1 \times E_1 \text{ or } (x, y) \in E_2 \times E_2, \\ 1 & \text{if } (x, y) \in E_1 \times E_2 \text{ or } (x, y) \in E_2 \times E_1. \end{cases}$$

(E, d) is separable metric. $\rho((v, i), (w, j)) := |v - w|$ is continuous and pos. definite.

We construct a Markov kernel P with at least 2 ipm and, for each $x, y \in E$, a coupling s.t. $\rho(X_n, Y_n) \rightarrow 0$ a.s.

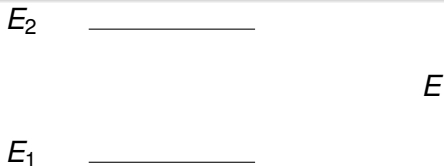
A counterexample

E_2 _____

E

E_1 _____

A counterexample



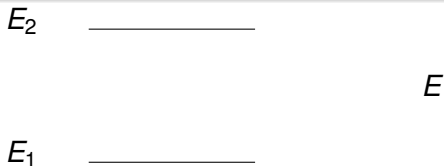
Construction of P

$$\pi_1(A) := \{z \in I : (z, 1) \in A\},$$

$$\pi_2(A) := \{z \in J : (z, 2) \in A\}, \quad A \subset E.$$

$$\mu(A) = \lambda^*(\pi_1(A)), \quad \nu(A) = \lambda^*(\pi_2(A)), \quad A \in \mathcal{B}(E)$$

A counterexample



Construction of P

$$\pi_1(A) := \{z \in I : (z, 1) \in A\},$$

$$\pi_2(A) := \{z \in J : (z, 2) \in A\}, \quad A \subset E.$$

$$\mu(A) = \lambda^*(\pi_1(A)), \quad \nu(A) = \lambda^*(\pi_2(A)), \quad A \in \mathcal{B}(E)$$

$$P(x, \cdot) = \begin{cases} \mu, & x \in E_1, \\ \nu, & x \in E_2. \end{cases}$$

A counterexample

E_2 _____

E

E_1 _____

Construction of P

$$\pi_1(A) := \{z \in I : (z, 1) \in A\},$$

$$\pi_2(A) := \{z \in J : (z, 2) \in A\}, \quad A \subset E.$$

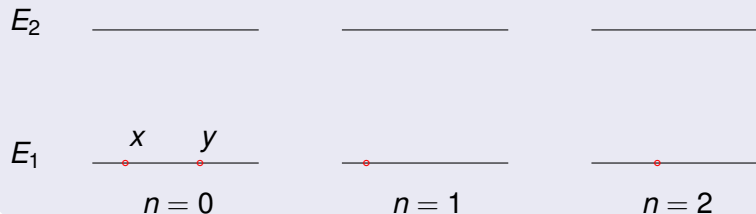
$$\mu(A) = \lambda^*(\pi_1(A)), \quad \nu(A) = \lambda^*(\pi_2(A)), \quad A \in \mathcal{B}(E)$$

$$P(x, \cdot) = \begin{cases} \mu, & x \in E_1, \\ \nu, & x \in E_2. \end{cases}$$

NB: μ and ν are mutually singular ipm. For $x, y \in E$ there is a coupling such that $\rho(X_n, Y_n) \rightarrow 0$.

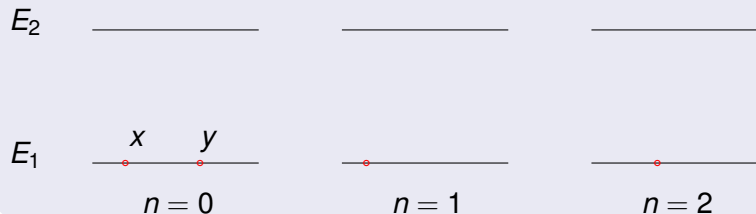
Construction of the coupling

First case: $x, y \in E_1$

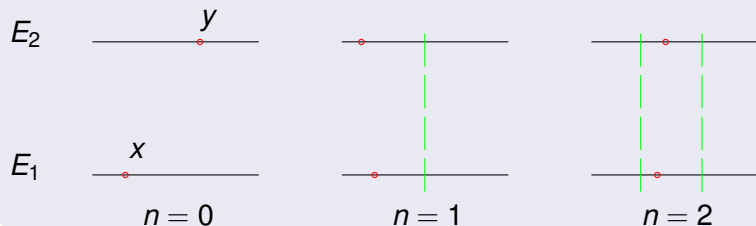


Construction of the coupling

First case: $x, y \in E_1$



Second case: $x \in E_1, y \in E_2$



Show uniqueness of weak solution of

$$dX(t) = f(X_t) dt + g(X_t) dW(t),$$

Show uniqueness of weak solution of

$$dX(t) = f(X_t) dt + g(X_t) dW(t), \quad X_0 = x \in \mathcal{C} = \mathcal{C}([-1, 0], \mathbb{R}^d)$$

with Hölder cont. f, g and nondegenerate g .

Show uniqueness of weak solution of

$$dX(t) = f(X_t) dt + g(X_t) dW(t), \quad X_0 = x \in \mathcal{C} = \mathcal{C}([-1, 0], \mathbb{R}^d)$$

with Hölder cont. f, g and nondegenerate g . Let

$$dY^\varepsilon(t) = f^\varepsilon(Y_t^\varepsilon) dt + g^\varepsilon(Y_t^\varepsilon) dW(t), \quad Y_0^\varepsilon = x$$

$$dZ^\varepsilon(t) = f^\varepsilon(Z_t^\varepsilon) dt + g^\varepsilon(Z_t^\varepsilon) dW(t) + \kappa_\varepsilon(X(t) - Z^\varepsilon(t)) dt, \quad Z_0^\varepsilon = x.$$

Then show for $T > 0$

- $\sup_{0 \leq t \leq T} |X(t) - Z^\varepsilon(t)| \rightarrow 0$ in probability,
- $d_{TV}(\mathcal{L}(Z^\varepsilon), \mathcal{L}(Y^\varepsilon)) \rightarrow 0$.

Show uniqueness of weak solution of

$$dX(t) = f(X_t) dt + g(X_t) dW(t), \quad X_0 = x \in \mathcal{C} = \mathcal{C}([-1, 0], \mathbb{R}^d)$$

with Hölder cont. f, g and nondegenerate g . Let

$$dY^\varepsilon(t) = f^\varepsilon(Y_t^\varepsilon) dt + g^\varepsilon(Y_t^\varepsilon) dW(t), \quad Y_0^\varepsilon = x$$

$$dZ^\varepsilon(t) = f^\varepsilon(Z_t^\varepsilon) dt + g^\varepsilon(Z_t^\varepsilon) dW(t) + \kappa_\varepsilon (X(t) - Z^\varepsilon(t)) dt, \quad Z_0^\varepsilon = x.$$

Then show for $T > 0$

- $\sup_{0 \leq t \leq T} |X(t) - Z^\varepsilon(t)| \rightarrow 0$ in probability,
- $d_{TV}(\mathcal{L}(Z^\varepsilon), \mathcal{L}(Y^\varepsilon)) \rightarrow 0$.

This implies weak uniqueness.

- S.: Couplings, gen. couplings and uniqueness of inv. measures, ECP, 2020
- Hairer, Mattingly, S.: Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations, PTRF, 2011
- Kulik, S.: Generalized couplings and convergence of transition probabilities, PTRF, 2018
- Kulik, S.: Well-posedness, stability, and sensitivities for stochastic delay equations: A generalized coupling approach, Ann. Probab., 2020
- S., Schindler: Convergence of Markov chain transition probabilities, ECP, 2021
- Butkovsky, Kulik, S.: Generalized couplings and ergodic rates for SPDEs and other Markov models, AAP, 2019

Thanks for your attention !