

Consider the following stochastic differential equation (SDE)

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, \quad (1)$$

where  $b$  is drift term,  $\sigma$  is the dispersion, and  $W_t$  is a one-dimensional Brownian motion. Here we consider functions  $b$  and  $\sigma$  to be Borel-measurable functions from  $\mathbb{R}_+ \times \mathbb{R}$  to  $\mathbb{R}$ . The SDE (1) can be equivalently written in the following integral form:

$$X_t = x + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s.$$

In this section, we are interested in the conditions under which the SDE (1) has strong existence and strong uniqueness. Recall that for an ordinary differential equation (ODE), shown as below:

$$x'(t) = b(t, x(t)), \quad (2)$$

the Picard-Lindelof/Cauchy-Lipschitz theorem states that if the function  $b$  is Lipschitz with respect to the state variable, then there exists a solution to the ODE (2). This is to say that there exists a constant  $K$  such that  $|b(t, x) - b(t, y)| \leq K|x - y|$ . For uniqueness of the solution, locally Lipschitz condition is enough. This is to say that for all  $R > 0$ , there exists a constant  $K_R$ , depending on  $R$ , such that  $|b(t, x) - b(t, y)| \leq K_R|x - y|$  for all  $x, y \in [-R, R]$ .

We can define similar conditions to show the strong existence and strong uniqueness for the SDE (1). We start by introducing the Gronwall inequality, which will be useful for proving the uniqueness results.

**Lemma 1. (Gronwall Inequality)**

Suppose that the measurable function  $f : [0, \infty) \rightarrow \mathbb{R}$  satisfies

$$0 \leq f(t) \leq a + b \int_0^t f(s)ds,$$

for some  $a, b \in \mathbb{R}$ . Then,

$$f(t) \leq ae^{bt}.$$

The proof for this Lemma is omitted, and we introduce the uniqueness result here.

**Theorem 1. (Strong Uniqueness)**

If for all  $R > 0$ , there exists a constant  $C_R$  such that for all  $t \geq 0$  and for all  $x, y \in [-R, R]$ , the following conditions hold for the functions  $b$  and  $\sigma$  in the SDE (1)

$$\begin{aligned} |b(t, x) - b(t, y)| &\leq C_R|x - y| \\ |\sigma(t, x) - \sigma(t, y)| &\leq C_R|x - y|, \end{aligned}$$

then strong uniqueness holds for the SDE (1).

*Proof.* Let  $X$  and  $\tilde{X}$  be two strong solutions to the SDE (1). Define  $Y_t := X_t - \tilde{X}_t$ . We would like to compute  $\mathbb{E}[Y_t^2]$ . From Ito's formula, we have:

$$\begin{aligned} Y_t^2 &= Y_0^2 + 2 \int_0^t Y_s dY_s + \int_0^t d\langle Y \rangle_s \\ &= 2 \int_0^t Y_s(b(s, X_s) - b(s, \tilde{X}_s))ds + 2 \int_0^t Y_s(\sigma(s, X_s) - \sigma(s, \tilde{X}_s))dW_s + \int_0^t (\sigma(s, X_s) - \sigma(s, \tilde{X}_s))^2 ds. \end{aligned} \quad (3)$$

Now, for a fixed  $R > 0$ , define the stopping time as follows

$$\tau_R = \inf\{t \geq 0 : |X_s| \geq R \text{ or } |\tilde{X}_s| \geq R\}.$$

With this localization, we would like to show that the stochastic integral is a martingale. We first show that the integrand is square-integrable.

$$\mathbb{E} \left[ \int_0^{t \wedge \tau_R} (Y_s(\sigma(s, X_s) - \sigma(s, \tilde{X}_s)))^2 ds \right] \leq C_R^2 \mathbb{E} \left[ \int_0^{t \wedge \tau_R} ((X_s - \tilde{X}_s))^4 ds \right] \leq 16C_R^2 R^2 T < \infty.$$

Thus, we know  $Y_s(\sigma(s, X_s) - \sigma(s, \tilde{X}_s)) \in \mathcal{L}_W^*$ , and the stochastic integral in (3) is (locally) a martingale, which implies that its expectation is zero. Then, we (locally) bound the expectation of (3) as follows:

$$\begin{aligned}
0 \leq \mathbb{E}[Y_{t \wedge \tau_R}^2] &= 2\mathbb{E} \left[ \int_0^{t \wedge \tau_R} Y_s(b(s, X_s) - b(s, \tilde{X}_s))ds \right] + \mathbb{E} \left[ \int_0^{t \wedge \tau_R} (\sigma(s, X_s) - \sigma(s, \tilde{X}_s))^2 ds \right] \\
&\leq 2C_R^2 \mathbb{E} \left[ \int_0^{t \wedge \tau_R} (X_s - \tilde{X}_s)^2 ds \right] + C_R^2 \mathbb{E} \left[ \int_0^{t \wedge \tau_R} (X_s - \tilde{X}_s)^2 ds \right] \\
&= \tilde{C} \mathbb{E} \left[ \int_0^t (X_{s \wedge \tau_R} - \tilde{X}_{s \wedge \tau_R})^2 ds \right] \\
&= \tilde{C} \int_0^t \mathbb{E} \left[ (X_{s \wedge \tau_R} - \tilde{X}_{s \wedge \tau_R})^2 \right] ds.
\end{aligned}$$

Now, we are ready to use Lemma 1. In particular, we have  $a = 0$ , which implies  $\mathbb{E}[Y_{t \wedge \tau_R}^2] = 0$ . Note  $R \rightarrow \infty \Rightarrow \tau_R \rightarrow \infty \Rightarrow t \wedge \tau_R = t$ . Then, by the dominated convergence theorem, we have  $\mathbb{E}[Y_t^2] = 0$ . This shows that  $\{X_t, 0 \leq t < \infty\}$  and  $\{\tilde{X}_t, 0 \leq t < \infty\}$  are modifications of each other. Since both solutions are continuous, we know that they are indistinguishable.  $\square$

Next, we show the existence of the strong solution to the SDE (1), which asks more than locally Lipschitz for the function  $b$  and  $\sigma$ .

**Theorem 2. (Strong Existence)**

Assume functions  $b$  and  $\sigma$  in the SDE (1) are Lipschitz continuous and satisfy the growth condition as detailed below:

$$\begin{aligned}
|b(t, x) - b(t, y)| &\leq C|x - y| \\
|\sigma(t, x) - \sigma(t, y)| &\leq C|x - y| \\
|b(t, x)|^2 + |\sigma(t, x)|^2 &\leq C(1 + |x|^2),
\end{aligned}$$

for some constant  $C$ , then the strong existence holds for the SDE (1).

Moreover, for every finite  $T > 0$ , there exists a positive constant  $\tilde{C}$ , depending only on  $T$  and  $C$  such that

$$\mathbb{E}[|X_t|^2] \leq \tilde{C}(1 + |x|^2)e^{\tilde{C}t}, \quad t \in [0, T]. \quad (4)$$

We first present the following lemma without prove.

**Lemma 2.** Consider the following sequence of stochastic processes  $\{X^{(k)}\}_{k \in \mathbb{N}}$ . Starting from  $X_t^{(0)} := x$ , define  $X^{(k)}$  recursively as:

$$X_t^{(k+1)} := x + \int_0^t b(s, X^{(k)})ds + \int_0^t \sigma(s, X^{(k)})dW_s. \quad (5)$$

Then, for each  $k \in \mathbb{N}$ , the process  $X_t^{(k)}$  is square integrable in the following sense: for every finite  $T > 0$ , there exists a positive constant  $\tilde{C}$ , depending only on  $T$  and  $C$  such that

$$\mathbb{E}[|X_t^{(k)}|^2] \leq \tilde{C}(1 + |x|^2)e^{\tilde{C}t}, \quad t \in [0, T], k \in \mathbb{N}. \quad (6)$$

Now, we prove Theorem 2.

*Proof.* The goal is to show that the sequence of stochastic processes  $\{X^{(k)}\}_{k \in \mathbb{N}}$  defined recursively in (5) converges to  $X_t$  so that the limit of (5) coincides with the SDE (1). Let's start with the following:

$$\mathbb{E} \left[ \max_{t \in [0, T]} |X_t^{(k+1)} - X_t^{(k)}|^2 \right] = \mathbb{E} \left[ \max_{t \in [0, T]} (B_t + M_t)^2 \right] \leq 2\mathbb{E} \left[ \max_{t \in [0, T]} B_t^2 \right] + 2\mathbb{E} \left[ \max_{t \in [0, T]} M_t^2 \right],$$

where

$$\begin{aligned}
B_t &:= \int_0^t b(s, X_s^{(k)}) - b(s, X_s^{(k-1)})ds \\
M_t &:= \int_0^t \sigma(s, X_s^{(k)}) - \sigma(s, X_s^{(k-1)})dW_s.
\end{aligned}$$

Let's first focus on the term  $B_t$ . Using the Cauchy-Schwarz inequality and the Lipschitz condition on the function  $b$ , we have:

$$\begin{aligned}
\mathbb{E} \left[ \max_{t \in [0, T]} B_t^2 \right] &= \mathbb{E} \left[ \max_{t \in [0, T]} \left| \int_0^t b(s, X_s^{(k)}) - b(s, X_s^{(k-1)}) ds \right|^2 \right] \\
&\leq \mathbb{E} \left[ \max_{t \in [0, T]} t \int_0^t |b(s, X_s^{(k)}) - b(s, X_s^{(k-1)})|^2 ds \right] \\
&= \mathbb{E} \left[ T \int_0^T |b(s, X_s^{(k)}) - b(s, X_s^{(k-1)})|^2 ds \right] \\
&\leq TC^2 \mathbb{E} \left[ \int_0^T |X_s^{(k)} - X_s^{(k-1)}|^2 ds \right].
\end{aligned}$$

Next, we turn attention to  $M_t$ . From Lemma 2 and the Lipschitz condition on the function  $\sigma$ , it is easy to show that  $M_t$  is a square integrable martingale. This fact allows us to use the **BDG-inequality**, which yields the following result

$$\begin{aligned}
\mathbb{E} \left[ \max_{t \in [0, T]} M_t^2 \right] &\leq \tilde{C} \mathbb{E}[\langle M \rangle_T] \\
&= \tilde{C} \mathbb{E} \left[ \int_0^T (\sigma(s, X_s^{(k)}) - \sigma(s, X_s^{(k-1)}))^2 ds \right] \\
&\leq \tilde{C} C^2 \mathbb{E} \left[ \int_0^T (X_s^{(k)} - X_s^{(k-1)})^2 ds \right].
\end{aligned}$$

Combining these two results, we have:

$$\begin{aligned}
\mathbb{E} \left[ \max_{t \in [0, T]} |X_t^{(k+1)} - X_t^{(k)}|^2 \right] &\leq 2C^2(T + \tilde{C}) \mathbb{E} \left[ \int_0^T |X_s^{(k)} - X_s^{(k-1)}|^2 ds \right] \\
&= L \mathbb{E} \left[ \int_0^T |X_s^{(k)} - X_s^{(k-1)}|^2 ds \right] \\
&\leq C^* \frac{(Lt)^k}{k!},
\end{aligned} \tag{7}$$

where

$$C^* = \mathbb{E} \left[ \max_{t \in [0, T]} |X_t^{(1)} - x|^2 \right].$$

The last inequality in (7) can be shown by induction. When  $k = 0$ , the result is obvious. Assuming  $k = k$  holds,  $k = k + 1$  can be verified by using the second last relation in (7).

From Markov's inequality, we have:

$$\mathbb{P} \left[ \max_{t \in [0, T]} |X_t^{(k+1)} - X_t^{(k)}| \geq 2^{-k} \right] \leq 4^k \mathbb{E} \left[ \max_{t \in [0, T]} |X_t^{(k+1)} - X_t^{(k)}|^2 \right] \leq C^* \frac{(4Lt)^k}{k!} \rightarrow 0,$$

which implies that  $\mathbb{P} \left[ \max_{t \in [0, T]} |X_t^{(k+1)} - X_t^{(k)}| \geq 2^{-k} \right]$  is a convergent series. From the Borel-Cantelli Lemma, we have:

$$\sum_{k=1}^{\infty} \mathbb{P} \left[ \max_{t \in [0, T]} |X_t^{(k+1)} - X_t^{(k)}| \geq 2^{-k} \right] < \infty \Rightarrow \mathbb{P} \left[ \limsup_{k \rightarrow \infty} \max_{t \in [0, T]} |X_t^{(k+1)} - X_t^{(k)}| \geq 2^{-k} \right] = 0.$$

In other words, eventually with probability 1,  $\max_{t \in [0, T]} |X_t^{(k+1)} - X_t^{(k)}| < 2^{-k}$ . It means that, with probability 1, the sequence  $\{X_t^{(k)}\}_{k \in \mathbb{N}}$  is a Cauchy sequence in  $\mathcal{C}([0, T], \mathbb{R})$ . Under the supremum norm (which makes the space complete), the sequence  $\{X_t^{(k)}\}_{k \in \mathbb{N}}$  converges.

Now, it remains to show that the continuous limit of  $\{X_t^{(k)}\}_{k \in \mathbb{N}}$ , denoted as  $X_t^\infty$ , is indeed the strong solution of the SDE (1). First of all, the integrability condition (4) can be shown from (6) plus the Fatou's Lemma. Next, it is obvious that  $X_t^\infty$  is  $\mathbb{F}$ -adapted and the initial condition holds. From the growth condition and (4), we know that  $b$  and  $\sigma^2$  are almost surely integrable. It remains to check the following:

$$\begin{aligned}\int_0^t b(s, X_s^{(k)}) ds &\rightarrow \int_0^t b(s, X_s^\infty) ds \\ \int_0^t \sigma(s, X_s^{(k)}) dW_s &\rightarrow \int_0^t \sigma(s, X_s^\infty) dW_s.\end{aligned}$$

The first convergence relation holds because  $b$  is Lipschitz continuous ( $b$  is uniformly bounded, and  $\{X_t^{(k)}(\omega)\}_{k \in \mathbb{N}}$  converges uniformly). The second convergence result will be shown using the [DDS theorem](#). Define  $\tilde{M}_t := \int_0^t \sigma(s, X_s^{(k)}) - \sigma(s, X_s^\infty) dW_s$ . It is easy to see that  $\tilde{M}_t \in \mathcal{M}_c^2$ , and  $\langle \tilde{M} \rangle_t \rightarrow \infty$ . Then, the martingale  $\tilde{M}_t$  is a time changed Brownian motion, i.e.,  $\tilde{M}_t = W_{\langle \tilde{M} \rangle_t}$ . Additionally, we know:  $\langle \tilde{M} \rangle_t = \int_0^t (\sigma(s, X_s^{(k)}) - \sigma(s, X_s^\infty))^2 ds = 0$ , because of the Lipschitz continuity of  $\sigma$  and the uniform convergence of  $\{X_t^{(k)}(\omega)\}_{k \in \mathbb{N}}$ . Thus, we have  $\tilde{M}_t = W_0 = 0$ , by definition of a standard Brownian motion. The proof is completed.  $\square$

Theorems 1 and 2 shows general conditions for strong existence and uniqueness for the SDE (1), and the proofs shown above can be generalized to multi-dimensional cases. However, the (locally) Lipschitz conditions are quite strong. Consider the following two SDEs.

**Example 1.** (*Squared-Bessel Process*)

$$dR_t = mdt + 2\sqrt{R_t}dW_t$$

**Example 2.** (*Squared Ornstein Uhlenbeck Processe*)

$$dR_t = a(m - R_t)dt + 2\sqrt{R_t}dW_t$$

We may show that the dispersion coefficients in both SDEs are not (locally) Lipschitz. We have shown before that the squared-Bessel process has a weak solution, but we cannot say more. If we can somehow show that the strong uniqueness holds, from a later study about weak solution, it also implies the weak uniqueness (in distribution). This is indeed true! We will see next that in the one-dimensional case, the Lipschitz condition on the dispersion coefficient can be relaxed considerably.

**Theorem 3.** (*Yamada & Watanabe (1971)*)

Suppose that the coefficients of the one dimensional SDE in (1) satisfy the conditions

$$\begin{aligned}|b(t, x) - b(t, y)| &\leq C|x - y| \\ |\sigma(t, x) - \sigma(t, y)| &\leq h(|x - y|),\end{aligned}\tag{8}$$

for every  $0 \leq t \leq \infty$  and  $x \in \mathbb{R}, y \in \mathbb{R}$ , where  $K$  is a positive constant and  $h : [0, \infty) \rightarrow [0, \infty)$  is a strictly increasing function with  $h(0) = 0$  and

$$\int_0^\epsilon h^{-2}(u)du = \infty, \quad \forall \epsilon > 0.\tag{9}$$

Then, strong uniqueness holds for the SDE (1).

*Proof.* Suppose  $X_t$  and  $\tilde{X}_t$  are two solutions to the SDE (1). We would like to show  $\mathbb{E}[|X_t - \tilde{X}_t|] = 0$ , which implies that  $X_t$  and  $\tilde{X}_t$  are modifications to each other, then by continuity,  $X_t$  and  $\tilde{X}_t$  are indistinguishable.

Unfortunately, the function  $|\cdot|$  is not twice differentiable, so that we cannot directly apply Ito's formula to it, like we did in the proof for Theorem 1. Instead, we construct a sequence of twice differentiable functions to approximate  $|\cdot|$ . From condition (9), we can construct a decreasing sequence  $\{a_n\}_{n=0}^\infty$  with  $a_0 = 1$ ,  $\lim_{n \rightarrow \infty} a_n = 0$ , and  $\int_{a_{n+1}}^{a_n} h^{-2}(u)du = n$ . Then, we construct a sequence of continuous functions  $\{\phi_n\}_{n=0}^\infty$  on  $\mathbb{R}$  with support in  $(a_{n+1}, a_n)$ , so that  $0 \leq \phi_n(x) \leq \frac{2}{nh^2(x)}$  and  $\int_{a_{n+1}}^{a_n} \phi_n(x)dx = 1$ . Now, we define the function  $\psi_n(x)$  as

$$\psi_n(x) := \int_0^{|x|} \int_0^y \phi_n(u)du dy.$$

$\psi_n(x)$  has the following properties:

1.  $\psi_n(x) = \psi_n(-x)$
2.  $|\psi'_n(x)| = |\int_0^{|x|} \phi_n(u)du| \leq 1$   
 $\psi''_n(x) = \phi_n(x)$   
 They implies that  $\psi_n(x)$  is twice continuously differentiable.
3.  $\psi_n(x) = |x|, \forall |x| \geq 1$

---


$$4. \lim_{n \rightarrow \infty} \psi_n(x) = |x|$$

Now, we apply Ito's formula as follows:

$$\begin{aligned} d\psi_n(X_t - \tilde{X}_t) &= \psi'_n(X_t - \tilde{x}_t)d(X_t - \tilde{x}_t) + \frac{1}{2}\psi''_n(X_t - \tilde{x}_t)d\langle X_t - \tilde{x}_t \rangle_t \\ &= \psi'_n(X_t - \tilde{x}_t)(b(t, X_t) - b(t, \tilde{X}_t))dt + \\ &\quad \psi'_n(X_t - \tilde{x}_t)(\sigma(t, X_t) - \sigma(t, \tilde{X}_t))dW_t + \\ &\quad \frac{1}{2}\psi''_n(X_t - \tilde{x}_t)(\sigma(t, X_t) - \sigma(t, \tilde{X}_t))^2dt. \end{aligned}$$

We claim  $\psi'_n(X_t - \tilde{x}_t)(\sigma(t, X_t) - \sigma(t, \tilde{X}_t)) \in \mathcal{L}_W^*$ , which can be shown using the localization argument. Then,  $\int_0^t \psi'_n(X_s - \tilde{x}_s)(\sigma(s, X_s) - \sigma(s, \tilde{X}_s))dW_s \in \mathcal{M}_c^2$ , so its expectation vanishes. We have the following

$$\begin{aligned} \mathbb{E}[\psi_n(X_t - \tilde{X}_t)] &= \mathbb{E}\left[\int_0^t \psi'_n(X_s - \tilde{x}_s)(b(s, X_s) - b(s, \tilde{X}_s))ds\right] + \\ &\quad \frac{1}{2}\mathbb{E}\left[\int_0^t \psi''_n(X_s - \tilde{x}_s)(\sigma(s, X_s) - \sigma(s, \tilde{X}_s))ds\right] \\ &\leq C\mathbb{E}\left[\int_0^t |X_s - \tilde{X}_s|ds\right] + \frac{1}{2}\mathbb{E}\left[\int_0^t \frac{2}{nh^2(|X_s - \tilde{X}_s|)}h^2(|X_s - \tilde{X}_s|)ds\right] \\ &= C\int_0^t \mathbb{E}[|X_s - \tilde{X}_s|]ds + \mathbb{E}\left[\int_0^t \frac{1}{n}ds\right]. \end{aligned}$$

Taking limit, we have

$$\mathbb{E}[|X_t - \tilde{X}_t|] \leq C \int_0^t \mathbb{E}[|X_s - \tilde{X}_s|]ds.$$

Applying the Gronwall inequality (1), we conclude that  $\mathbb{E}[|X_t - \tilde{X}_t|] = 0$ .

□

Sometimes the drift term in the SDE (1) can be too complicated to analyze. If we can either upper or lower bound it by another function that is easier to analyze, then we can compare the solutions of these two SDEs as shown in the following results. Note, this result also only works for one-dimensional SDEs.

**Theorem 4. (Comparison Principle)**

Let  $X_t$  and  $\tilde{X}_t$  be strong solutions to the SDE (1) with drift term  $b(t, x)$  and  $\tilde{b}(t, x)$ , respectively. Suppose the following conditions hold:

1.  $b, \tilde{b}$ , and  $\sigma$  are continuous functions
2.  $|\sigma(t, x) - \sigma(t, y)| \leq h(|x - y|)$ , with  $h$  being the function in Theorem 3
3.  $b(t, x)$  or  $\tilde{b}(t, x)$  satisfies the Lipschitz condition.
4. The initial condition  $x \leq \tilde{x}$ .
5.  $b(t, x) \leq \tilde{b}(t, x)$  for all  $t$  and  $x$

Then,  $X_t \leq \tilde{X}_t$  for all  $t$ .

*Proof.* The proof is similar to the previous one, except we would like to show  $\mathbb{E}[|X_t - \tilde{X}_t|_+] := \mathbb{E}[\max\{|X_t - \tilde{X}_t|, 0\}] = 0$ . Define  $\psi_n(x) = \tilde{\psi}_n(x)\mathbb{1}_{x \geq 0}$ , where  $\tilde{\psi}_n(x)$  is the same twice continuously differentiable function in the proof for Theorem 3. Without loss of generality, we assume that  $\tilde{b}(t, x)$  is Lipschitz continuous. We omit first several steps related to the Ito calculus, because they are identical to the previous proof. Taking expectation, we have

$$\begin{aligned}
\mathbb{E}[\psi_n(X_t - \tilde{X}_t)] &= \mathbb{E}\left[\int_0^t \psi'_n(X_s - \tilde{x}_s)(b(s, X_s) - \tilde{b}(s, \tilde{X}_s))ds\right] + \frac{1}{2}\mathbb{E}\left[\int_0^t \psi''_n(X_s - \tilde{x}_s)(\sigma(s, X_s) - \sigma(s, \tilde{X}_s))ds\right] \\
&= \mathbb{E}\left[\int_0^t \psi'_n(X_s - \tilde{x}_s)(b(s, X_s) - \tilde{b}(s, X_s))ds\right] + \mathbb{E}\left[\int_0^t \psi'_n(X_s - \tilde{x}_s)(\tilde{b}(s, X_s) - \tilde{b}(s, \tilde{X}_s))ds\right] + \\
&\quad \frac{1}{2}\mathbb{E}\left[\int_0^t \psi''_n(X_s - \tilde{x}_s)(\sigma(s, X_s) - \sigma(s, \tilde{X}_s))ds\right] \\
&\leq \mathbb{E}\left[\int_0^t \psi'_n(X_s - \tilde{x}_s)(\tilde{b}(s, X_s) - \tilde{b}(s, \tilde{X}_s))ds\right] + \frac{1}{2}\mathbb{E}\left[\int_0^t \psi''_n(X_s - \tilde{x}_s)(\sigma(s, X_s) - \sigma(s, \tilde{X}_s))ds\right] \\
&\leq C\mathbb{E}\left[\int_0^t |X_s - \tilde{X}_s|ds\right] + \frac{1}{2}\mathbb{E}\left[\int_0^t \frac{2}{nh^2(|X_s - \tilde{X}_s|)}h^2(|X_s - \tilde{X}_s|)ds\right] \\
&= C\int_0^t \mathbb{E}[|X_s - \tilde{X}_s|]ds + \mathbb{E}\left[\int_0^t \frac{1}{n}ds\right].
\end{aligned}$$

Then, using the Gronwall inequality, we have  $\mathbb{E}[|X_t - \tilde{X}_t|_+] = 0$ , which implies  $X_t \leq \tilde{X}_t$  for all  $t$  almost surely.  $\square$

Now, we shift gear to look at the weak solution of an SDE. Consider the following example.

**Example 3. Tanaka's SDE**

Consider the following SDE

$$dX_t = \text{sign}\{X_t\}dW_t, \quad X_0 = 0$$

It can be shown that this SDE does not have a strong solution, but we take it for granted. Suppose that a weak solution  $X_t$  exists. It is clear that  $X_t \in \mathcal{M}_c^2$ . From below

$$\langle X \rangle_t = \int_0^t (\text{sign}(X_t))^2 dt = t,$$

we may use Levy's characterization to conclude that  $X_t$  is a Brownian motion. Additionally, it is true for every weak solution, so we have weak uniqueness. However, notice that if  $X_t$  is a solution, then  $-X_t$  is also a solution, which implies that strong uniqueness does not hold.

Now let's show that weak solution indeed exists. Consider the augmented filtered probability space  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{\mathbb{P}})$ , carrying a Brownian motion  $\tilde{X}_t$ . Our goal is to show that  $\tilde{X}_t$  satisfies the SDE. Define  $\tilde{W}_t := \int_0^t \text{sign}(\tilde{X}_s)d\tilde{X}_s$ . It is easy to show that  $\tilde{W}_t$  is a Brownian motion through Levy's characterization. Then

$$\int_0^t \text{sign}(\tilde{X}_s)d\tilde{W}_s = \int_0^t \text{sign}(\tilde{X}_s)\text{sign}(\tilde{X}_s)d\tilde{x}_s = \int_0^t d\tilde{X}_s = \tilde{X}_t.$$

We showed that  $\tilde{X}_t$  is indeed a weak solution.

Now, we present the conditions that guarantee weak existence and uniqueness for a special form of the SDE (1).

**Theorem 5.** Consider the following SDE

$$dX_t = b(t, X_t)dt + dW_t, \quad t \in [0, T], \quad X_0 = x \in \mathbb{R}. \quad (10)$$

If  $b$  is a bounded and measurable function, then weak existence and uniqueness hold for this SDE.

*Proof.* Consider the filtered probability space  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ , carrying a Brownian motion  $X_t$ . By the condition on  $b$  given in the assumption, we know  $\mathbb{E}[\exp(\frac{1}{2}\int_0^T |b(s, X_s)|^2 ds)] < \infty$ . Then, Nikonov's condition tells us that we can properly define a measure  $\mathbb{Q}$  as follows

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(\int_0^T b(s, X_s)dW_s - \frac{1}{2}\int_0^T |b(s, X_s)|^2 ds\right),$$

and

$$\tilde{W}_t := X_t - x - \int_0^t b(s, X_s)ds$$

is a Brownian motion under the measure  $\mathbb{Q}$ , and  $\mathbb{Q}[\tilde{W}_0 = 0] = 1$ . Rearranging the terms, we get

$$dX_t = b(t, X_t)dt + d\tilde{W}_t.$$

Thus, we have  $(X, \tilde{W}), (\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q})$  as the weak solution.

Next, we prove the weak uniqueness. Consider  $(X, W)$  and  $(\tilde{X}, \tilde{W})$  being two weak solutions, possible defined on different probability spaces. We need to show that they equal in distribution. Assume that  $(X, W), (\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q})$  is a weak solution to the SDE. With the same Nikonov's condition as above, we can properly define the measure  $\mathbb{Q}$  as follows

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(-\int_0^T b(s, X_s)dW_s - \frac{1}{2}\int_0^T |b(s, X_s)|^2 ds\right),$$

and

$$\tilde{W}_t := W_t + \int_0^t b(s, X_s)ds$$

is a Brownian motion under the measure  $\mathbb{Q}$ . Since  $X_t$  satisfies the SDE, we have

$$X_t := x + \tilde{W}_t.$$

For a measurable set  $A$ , we have

$$\begin{aligned} \mathbb{P}[\{X_t\}_{t \in [0, T]} \in A] &= \mathbb{E}^{\mathbb{P}}[\mathbb{1}_{\{X_t\}_{t \in [0, T]} \in A}] \\ &= \mathbb{E}^{\mathbb{Q}}\left[\frac{d\mathbb{P}}{d\mathbb{Q}} \mathbb{1}_{\{X_t\}_{t \in [0, T]} \in A}\right] \\ &= \mathbb{E}^{\mathbb{Q}}\left[\exp\left(\int_0^T b(s, X_s)dW_s + \frac{1}{2}\int_0^T |b(s, X_s)|^2 ds\right) \mathbb{1}_{\{X_t\}_{t \in [0, T]} \in A}\right] \\ &= \mathbb{E}^{\mathbb{Q}}\left[\exp\left(\int_0^T b(s, X_s)d\tilde{W}_s - \frac{1}{2}\int_0^T |b(s, X_s)|^2 ds\right) \mathbb{1}_{\{X_t\}_{t \in [0, T]} \in A}\right] \\ &= \mathbb{E}^{\mathbb{Q}}\left[\exp\left(\int_0^T b(s, X_s)dX_s + \frac{1}{2}\int_0^T |b(s, X_s)|^2 ds\right) \mathbb{1}_{\{X_t\}_{t \in [0, T]} \in A}\right], \end{aligned}$$

which is enough to show the weak uniqueness, because  $\mathbb{P}[\{X_t\}_{t \in [0, T]} \in A]$  is uniquely defined, giving that  $X_t$  is a Brownian motion under the measure  $\mathbb{Q}$ .  $\square$

**Remark 1.** This theorem represents the regularization by noise. For the ODE counterpart, Lipschitz condition is required for existence and uniqueness, which is a stronger condition on  $b$ .

**Remark 2.** More refined conditions for the function  $b$  in the SDE (10) that guarantees the weak existence and uniqueness existed.

**Remark 3.** This theorem can be extended to the infinite time horizon case with the Kolmogorov's extension theorem.

**Remark 4.** Clearly, strong existence implies weak existence, and strong uniqueness implies weak uniqueness if a strong solution exists. However, it is much more tricky to conclude that strong uniqueness implies weak uniqueness for weak solutions, which leads to the following theorem.

**Theorem 6.** Pathwise uniqueness implies uniqueness in the sense of probability law.

*Proof.* The proof is quite involved, so we only show the basic idea behind it. Consider two weak solutions to the SDE (10),  $(X, W)$  and  $(\tilde{X}, \tilde{W})$ . Since they possibly live in different probability spaces, we have to create a new canonical space and copy both solutions to the same probability space. Then, pathwise uniqueness in the new space directly implies uniqueness in distribution.  $\square$

The weak solution to an SDE discussed so far is obtained through finding a process with the right law (and doing so uniquely for weak uniqueness). A variation on this approach, developed by Stroock & Varadhan (1969), formulates the search for the law of a diffusion process with given drift and dispersion coefficients in terms of a martingale problem.

Consider the same SDE as in (1), which we copy to here for better reference

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t,$$

where  $X_t \in \mathbb{R}^d$ ,  $b : [0, \infty) \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ ,  $\sigma : [0, \infty) \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$ , and  $W$  is a  $d$ -dimensional Brownian motion. In other words, each component  $X^i$  satisfies the following dynamics:

$$dX_t^i = b_i(t, X_t)dt + \sum_{j=1}^d \sigma_{i,j}(t, X_t)dW_t^j. \quad (11)$$

The quadratic variations are given by

$$\begin{aligned} \langle X^i, X^j \rangle_t &= \left\langle \sum_{k=1}^d \int_0^t \sigma_{i,k}(s, X_s)dW_s^k, \sum_{l=1}^d \int_0^t \sigma_{j,l}(s, X_s)dW_s^l \right\rangle_t \\ &= \int_0^t \sum_{k=1}^d \sum_{l=1}^d \sigma_{i,k}(s, X_s)\sigma_{j,l}(s, X_s)d\langle W^k, W^l \rangle_s \\ &= \int_0^t \sum_{k=1}^d \sigma_{i,k}(s, X_s)\sigma_{j,k}(s, X_s)ds. \end{aligned}$$

Let  $f \in C^2(\mathbb{R}^d, \mathbb{R})$ . We first claim that if  $X$  is a solution to the SDE, then the process  $M^f$  defined by

$$M_t^f := f(X_t) - f(X_0) - \int_0^t (\mathcal{A}f)(s, X_s)ds, \quad (12)$$

where

$$(\mathcal{A}f)(t, x) := \sum_{i=1}^d b_i(t, x)\partial_{x^i}f(x) + \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d \left( \sum_{k=1}^d \sigma_{i,k}(t, x)\sigma_{j,k}(t, x) \right) \partial_{x^i, x^j}^2 f(x) \quad (13)$$

is a local martingale.

This claim is immediate after applying Ito's formula to the solution of the SDE as detailed below.

$$\begin{aligned} f(X_t) &= f(X_0) + \sum_{i=1}^d \int_0^t \partial_{x^i}f(X_s)dX_s^i + \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d \int_0^t \partial_{x^i, x^j}^2 f(X_s)d\langle X^i, X^j \rangle_s \\ &= f(X_0) + \sum_{i=1}^d \int_0^t \partial_{x^i}f(X_s)dX_s^i + \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d \int_0^t \sum_{k=1}^d \sigma_{i,k}(s, X_s)\sigma_{j,k}(s, X_s)\partial_{x^i, x^j}^2 f(X_s)ds \\ &= f(X_0) + \int_0^t (\mathcal{A}f)(s, X_s)ds + \int_0^t \partial_{x^i}f(X_s)b_i(s, X_s)ds + \sum_{i=1}^d \sum_{j=1}^d \int_0^t \partial_{x^i}f(X_s)\sigma_{i,j}(s, X_s)dW_s^j. \end{aligned}$$

By the definition of the operator  $\mathcal{A}$  in (13), we have

$$M_t^f = \sum_{i=1}^d \sum_{j=1}^d \int_0^t \partial_{x^i}f(X_s)\sigma_{i,j}(s, X_s)dW_s^j,$$

which is a sum of local martingales (because each stochastic integral is a local martingale with  $W_t \in \mathcal{M}_c^2$ ). Thus,  $M_t^f$  is a local martingale.

Now, we formally define the martingale problem.

**Definition 1. (Martingale Problem)**

A probability measure  $\mathbb{P}$  on  $(C([0, \infty)^d, \mathcal{B}(C[0, \infty)^d))$ , under which

$$M_t^f := f(X_t) - f(X_0) - \int_0^t (\mathcal{A}f)(s, X_s)ds, \quad (14)$$

is a continuous local martingale for every  $f \in C^2(\mathbb{R}^d)$ , is called a solution to the local martingale problem associated with  $\mathcal{A}$ .

What we have shown before is that a weak solution to the SDE induces on  $(C([0, \infty)^d, \mathcal{B}(C[0, \infty)^d))$  a probability measure  $\mathbb{P}$  which solves the local martingale problem associated with  $\mathcal{A}$ . Surprisingly, the converse is also true!

**Proposition 1.** Let  $\mathbb{P}$  be a probability measure on  $(C([0, \infty)^d, \mathcal{B}(C[0, \infty)^d))$  under which the process  $M_t^f$  of (14) is a continuous local martingale for the choices  $f(x) = x_i$  and  $f(x) = x_i x_j$ ;  $1 \leq i, j \leq d$ . Then, there is an  $r$ -dimensional Brownian motion  $W$  defined on an extension  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$  of  $(C([0, \infty)^d, \mathcal{B}(C[0, \infty)^d), \mathbb{P})$ , such that  $(X_t, W_t), (\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$  is a weak solution to the SDE.



*Proof.* Let  $i = 1, \dots, d$  and  $f(x) = x_i$ , we have that  $M_t^i$  defined by

$$M_t^i = X_t^i - X_0^i - \int_0^t b_i(s, X_s) ds \quad (15)$$

is a continuous local martingale by assumption. Now, let  $i, j = 1, \dots, d$  and  $f(x) = x_i x_j$ , we have that  $M_t^{i,j}$  defined by

$$M_t^{i,j} = X_t^i X_t^j - X_0^i X_0^j - \int_0^t b_i(s, X_s) + b_j(s, X_s) + a_{i,j}(s, X_s) ds,$$

where

$$a_{i,j}(s, X_s) = \sum_{k=1}^d \sigma_{i,k}(s, X_s) \sigma_{j,k}(s, X_s) ds$$

is a continuous local martingale by assumption. We may show that

$$M_t^i M_t^j - \int_0^t a_{i,j}(s, X) ds$$

can be expressed as a sum of continuous local martingale  $M_t^{i,j} - X_0^i M_t^j - X_0^j M_t^i$  and the process

$$\begin{aligned} & \int_0^t (X_s^i - X_t^i) b_j(s, X) ds + \int_0^t (X_s^j - X_t^j) b_i(s, X) ds + \int_0^t b_i(s, X) ds \int_0^t b_j(s, X) ds \\ &= \int_0^t (M_s^i - M_t^i) b_j(s, X) ds + \int_0^t (M_s^j - M_t^j) b_i(s, X) ds \\ &= - \int_0^t \int_0^s b_j(u, X) du dM_s^i - \int_0^t \int_0^s b_i(u, X) du dM_s^j, \end{aligned}$$

where the last inequality holds by using Ito's formula. In summary, we have

$$M_t^i M_t^j - \int_0^t a_{i,j}(s, X) ds = M_t^{i,j} - X_0^i M_t^j - X_0^j M_t^i - \int_0^t \int_0^s b_j(u, X) du dM_s^i - \int_0^t \int_0^s b_i(u, X) du dM_s^j$$

where the right hand side is a continuous local martingale, which implies the left hand side is also a continuous local martingale. Additionally, we have

$$\langle M^i, M^j \rangle_t = \int_0^t a_{i,j}(s, X) ds, \quad (16)$$

which implies that  $d\langle M^i, M^j \rangle_t \ll dt$ , and we can apply the Martingale Representation theorem. The theorem says that there exists a  $d$ -dimensional Brownian motion  $\tilde{W}$ , defined on an extension  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$  of  $(C([0, \infty)^d, \mathcal{B}(C[0, \infty)^d), \mathbb{P})$  as well as a matrix  $\rho$  of measurable adapted processes with

$$\tilde{\mathbb{P}} \left[ \int_0^t \rho_{i,j}^2(s) ds < \infty \right] = 1,$$

such that

$$M_t^i = \sum_{j=1}^d \int_0^t \rho_{i,j}(s) d\tilde{W}_s^j.$$

By (16), we must have  $\rho \rho^T = \sigma \sigma^T = a$ . By the original definition of  $M_t^i$  in (15), we have

$$X_t^i = M_t^i + X_0^i + \int_0^t b_i(s, X_s) ds = X_0^i + \int_0^t b_i(s, X_s) ds + \sum_{j=1}^d \int_0^t \rho_{i,j}(s) d\tilde{W}_s^j.$$

Comparing this identity to (11), if we can show that there exists a Brownian motion  $W$  on the extended probability space such that

$$\int_0^t \rho(s) d\tilde{W}_s = \int_0^t \sigma(s, X_s) dW_s,$$

then we are done. From the Lemma shown below, it turns out that we can indeed define the following process

$$W_t := \int_0^t R^T(\rho(s), \sigma(s, X)) d\tilde{W}_s,$$

which satisfies  $\sigma = \rho R(\rho, \sigma)$  and  $R(\rho, \sigma)R(\rho, \sigma)^T = I$ . Then,  $W_t$  is a continuous local martingale and

$$\langle W^i, W^j \rangle_t = t\delta_{i,j},$$

which, by Levy's characterization,  $W_t$  is a Brownian motion. The proof is complete.  $\square$

**Lemma 3.** *There exists a Borel-measurable,  $(d \times d)$ -matrix-valued function  $R(\rho, \sigma)$  defined on the set*

$$D := \{(\rho, \sigma) : \rho \text{ and } \sigma \text{ are } (d \times d) \text{ matrices with } \rho\rho^T = \sigma\sigma^T\}$$

*such that*

$$\sigma = \rho R(\rho, \sigma), \quad R(\rho, \sigma)R^T(\rho, \sigma) = I, \quad (\rho, \sigma) \in D.$$

*Proof.* Let  $A := \rho\rho^T = \sigma\sigma^T$ .  $A$  is symmetric and positive semidefinite, which allows us to diagonalize  $A$  by orthogonal matrix  $Q$  as  $Q A Q^T = \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix}$ , where  $\Lambda$  is a  $(k \times k)$  diagonal matrix with eigenvalue of  $A$  on the diagonal. Then, we have  $Q A Q^T = (Q\rho)(Q\rho)^T = (Q\sigma)(Q\sigma)^T$ , which implies  $Q\rho = \begin{bmatrix} Y_1 \\ 0 \end{bmatrix}$ , where  $Y_1 Y_1^T = \Lambda$ .

Similarly, we have  $Q\sigma = \begin{bmatrix} Z_1 \\ 0 \end{bmatrix}$ , where  $Z_1 Z_1^T = \Lambda$ .

We can find an orthonormal basis for the  $(d - k)$ -dimensional subspace orthogonal to the space spanned by the rows of  $Y_1$ . Let  $Y_2$  be the  $(d - k) \times d$  matrix consisting the orthonormal row vectors of this basis, and define  $Y := \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$ . We have  $Y Y^T = \begin{bmatrix} \Lambda & 0 \\ 0 & I \end{bmatrix}$ , and  $Q\rho Y^T = \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix}$ . Similarly, we may define  $Z := \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}$ , such that  $Z Z^T = \begin{bmatrix} \Lambda & 0 \\ 0 & I \end{bmatrix}$ , and  $Q\sigma Z^T = \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix}$ . Similarly, we may define  $Z := \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}$ .

Now, we define  $R := Y^T(Z^T)^{-1}$ . One may check that it satisfies all the requirements.  $\square$