Functional limit theorems for approximating irregular SDEs, general diffusions and their exit times

Mikhail Urusov
Universität Duisburg-Essen

Joint work with Stefan Ankirchner and Thomas Kruse

HSE / LSA winter meeting, Snegiri, Moscow region, 03.12.2019

Aims

- Approximate 1d continuous strong Markov processes
- Approximate exit times
- Establish convergence rates

1d continuous strong Markov processes

General diffusions, 1

General diffusion $\equiv$ 1d continuous strong Markov process

Throughout we consider a general diffusion $Y = (Y_t)_{t \geq 0}$

State space $\mathcal{I} = \langle \ell, r \rangle$, $-\infty \leq \ell < r \leq \infty$

- Notation: $\mathcal{I}^\circ = (\ell, r)$, $\overline{\mathcal{I}} = [\ell, r]$
- Y regular: $\forall y \in \mathcal{I}^\circ \; \forall x \in \mathcal{I}$: $\mathsf{P}_y(H_x < \infty) > 0$

$$H_x = \inf\{t \geq 0 : Y_t = x\}$$

General diffusions, 2

Law(Y) specified by

- scale function $s\colon \mathcal{I}^\circ \to \mathbb{R}$ $\quad (\nearrow\nearrow)$
- speed measure m on $\mathcal{I}$ with the property

$$0 < m([a,b]) < \infty \quad \forall a < b \text{ in } \mathcal{I}^\circ$$

W.l.o.g.: Y in natural scale, i.e., $s = id$

Input: speed measure m

For the moment assume that accessible boundary points are absorbing

Example 1: solutions of driftless SDEs

Solutions of SDEs ($\equiv$ diffusions) are general diffusions

In natural scale $\iff$ driftless

$$dY_t = \eta(Y_t)\,dW_t$$

Engelbert–Schmidt conditions (for weak $\exists$-ce and uniqueness in law):

$$\eta(x) \neq 0 \ \forall x \in \mathcal{I}^\circ$$
$$\eta^{-2} \in L^1_{loc}(\mathcal{I}^\circ)$$

Convention: ℓ, r absorbing whenever accessible

Speed measure: $m(dx) = \frac{2}{\eta^2(x)}\,dx \quad (\Longrightarrow m \ll \mu_L)$

Special case of a Brownian motion on $\mathbb{R}$: $m = 2\mu_L$

Why seeking for a new scheme?

The question to approximate Y arises even within Example 1

- Usual approximation schemes can fail when η is irregular
- E.g., the Euler scheme does not converge to Y (even weakly!) in the case

$$dY_t = \frac{1}{|Y_t|}\, dW_t$$

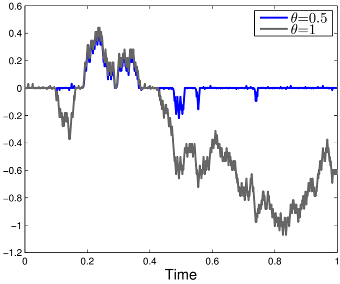
Even more interesting question:
how to approximate a general diffusion with no SDE behind?

Example 2: a general diffusion with no SDE behind

Violating $m \ll \mu_L$ provides examples of general diffusions that cannot be realized via SDEs

Example: BM on $\mathbb{R}$ sticky at 0

$$m(dx) = 2\,dx + \frac{2}{\theta}\,\delta_{\{0\}}(dx), \quad \theta \in (0, \infty)$$



Recent interest in “stickiness”

Some references (an incomplete list):

Amir 1991
Fushiya 2010
Karatzas–Shiryaev–Shkolnikov 2011
Bass 2014
Engelbert–Peskir 2014
Hajri–Caglar–Arnaudon 2017
Grothaus–Voßhall 2018
Eberle–Zimmer 2019
Konarovskiy–von Renesse 2020+

FLT on the path space

Preprint on arXiv:

- Ankirchner, Kruse, Urusov. “Functional limit theorem”, 2019

Euler and weak Euler schemes

Motivation via the (weak) Euler scheme for

$$dY_t = \eta(Y_t)\,dW_t$$

Time step: $h > 0$ (small)

$$X^{h,Eu}_{(k+1)h} = X^{h,Eu}_{kh} + \eta(X^{h,Eu}_{kh})\sqrt{h}\,\zeta_{k+1}$$

- Euler: (ζ_j) iid $\sim N(0,1)$
- weak Euler: (ζ_j) iid, $\mathrm{E}[\zeta_j^k] = \mathrm{E}[N(0,1)^k]$, $k = 1,2,3$

Class of approximation schemes

We consider the class of approximating schemes of the form

$$X^h_{(k+1)h} = X^h_{kh} + a_h(X^h_{kh})\,\xi_{k+1}$$

with symmetric Bernoulli (ξ_j): (ξ_j) iid, $\mathsf{P}(\xi_j = \pm 1) = \frac{1}{2}$

- To get weak Euler scheme with symmetric Bernoulli (ζ_j) take $a_h(y) = \eta(y)\sqrt{h}$
- But we do not require this, as there is no η in general

We refer to the function a_h as to a **scale factor:**
For each $h \in (0, \overline{h})$ let $a_h \colon \mathcal{I} \to [0, \infty)$ be such that

$$\forall y \in \mathcal{I} \colon\; y \pm a_h(y) \in \mathcal{I}$$

In particular, $a_h(\ell) = a_h(r) = 0$ (whenever ℓ and/or $r \in \mathcal{I}$)

$(X^h_{kh})_{k \in \mathbb{N}_0} \rightsquigarrow (X^h_t)_{t \geq 0}$ via linear interpolation

Convergence

Challenge: find conditions on m and $(a_h)_{h\in(0,\bar{h})}$ to ensure convergence of $(X^h)_{h\in(0,\bar{h})}$ to Y as $h \to 0$

Condition (A): for any compact subset $K \subset \mathcal{I}^\circ$,

$$\sup_{y\in K} \left| \frac{1}{2} \int_{(y-a_h(y),y+a_h(y))} (a_h(y) - |u-y|)\, m(du) - h \right| = o(h), \quad h \to 0$$

Theorem (FLT on the path space)

Under Condition (A),

$$\text{Law}(X_t^h; t \geq 0) \xrightarrow{w} \text{Law}(Y_t; t \geq 0).$$

Why it works

- For $y \in \mathcal{I}^\circ$ and $a > 0$ such that $y \pm a \in \mathcal{I}$

$$\mathsf{E}_y[H_{y-a,y+a}] = \frac{1}{2}\int_{(y-a,y+a)} (a - |u-y|)\, m(dy),$$

  where $H_{a,b} = H_a \wedge H_b$
- Hence, Condition (A) is nothing else but the requirement on the scale factors $(a_h)_{h\in(0,\overline{h})}$ to be such that

$$\sup_{y\in K} \left|\mathsf{E}_y[H_{y-a_h(y),y+a_h(y)}] - h\right| = o(h), \quad h \to 0 \qquad (*)$$

- (Main idea) As the approximation jumps from each point y into $y \pm a_h(y)$, we can construct stopping times (τ_k^h) such that

$$(Y_{\tau_k^h}; k \in \mathbb{N}_0) \overset{d}{=} (X_{kh}^h; k \in \mathbb{N}_0),$$

  while $(*)$ translates into $\mathsf{E}[\tau_{k+1}^h - \tau_k^h] \approx h \quad$ (informally)

EMCEL scheme (idea)

- For $y \in \mathcal{I}^\circ$, choose $a_h(y)$ so that

$$\mathsf{E}_y[H_{y-a_h(y),y+a_h(y)}] = h$$

- To this end, for each $y \in \mathcal{I}^\circ$, solve w.r.to $a > 0$

$$\frac{1}{2}\int_{(y-a,y+a)} (a - |u-y|)\, m(du) = h$$

- Solution: $\widehat{a}_h(y)$
  Associated Markov chain: $(\widehat{X}_t^h)_{t\geq 0}$
  $\longrightarrow$ EMCEL(h) scale factor / approximation

EMCEL(h): **E**mbeddable **M**arkov **C**hain with **E**xpected time **L**ag h

Intermediate summary

- The EMCEL scheme always satisfies Condition (A)
  $\Longrightarrow$ **every** general diffusion can be approximated by EMCEL
- In particular, no problem with
  $$dY_t = \frac{1}{|Y_t|}\, dW_t,$$
  no problem with stickiness, etc.
- In case when the equation determining EMCEL cannot be solved in closed form, Condition (A) dictates the required precision for the numerical solution in order to retain the convergence
- Alternatively, Condition (A) can be used to verify convergence for other schemes, e.g., for the weak Euler one

Application to the weak Euler scheme

- $dY_t = \eta(Y_t)\, dW_t$

Corollary

Assume $\mathcal{I} = \mathbb{R}$, η continuous and nonvanishing. Consider the weak Euler scheme $X^{h,Eu}$ (generated by the scale factors $a_h(y) = \eta(y)\sqrt{h}$). Then

$$\mathrm{Law}(X_t^{h,Eu}; t \geq 0) \xrightarrow{w} \mathrm{Law}(Y_t; t \geq 0).$$

- New result
- In a sense, this is a result between Gyöngy (1998) and Yan (2002) (although they consider the classical rather than weak Euler scheme)

FLT for processes & exit times

Preprint on arXiv:

- Kruse, Urusov. “Exit times”, 2019

General problem formulation

Recall the notation

$$H_b(Y) = \inf\{t \geq 0 : Y_t = b\}$$

Q: for $b \in \mathcal{I}$,

$$X^h \xrightarrow[h \searrow 0]{w} Y \quad \stackrel{?}{\Longrightarrow} \quad H_b(X^h) \xrightarrow[h \searrow 0]{w} H_b(Y)$$

A: for $b \in \mathcal{I}^\circ$, **true**, as the path functional H_b P_Y-a.s. continuous;
for $b \in \{\ell, r\}$, in general, **false** (see below)

- H_ℓ (resp. H_r) essentially discontinuous path functional whenever ℓ (resp. r) is an accessible boundary

We refer to $H_\ell(Y)$ and $H_r(Y)$ as to **exit times** of Y

Conditions (idea)

(Recall) **Condition (A):** for any compact subset $K \subset \mathcal{I}^\circ$,

$$\sup_{y\in K}\left|\frac{1}{2}\int_{(y-a_h(y),y+a_h(y))}(a_h(y)-|u-y|)\,m(du)-h\right|=o(h),\quad h\to 0$$

Condition (B): for the moment refrain from its formulation

Condition (D):

$$\sup_{y\in\mathcal{I}}\left|\frac{1}{2}\int_{(y-a_h(y),y+a_h(y))}(a_h(y)-|u-y|)\,m(du)-h\right|=o(h),\quad h\to 0$$

Relation: (D) $\Rightarrow$ (B) $\Rightarrow$ (A)

FLT for processes & exit times

(Recall) Theorem 1: Under Condition (A), $X^h \xrightarrow[h\searrow 0]{w} Y$

Theorem 2: Under Condition (B),

$$(X^h, H_\ell(X^h), H_r(X^h)) \xrightarrow[h\searrow 0]{w} (Y, H_\ell(Y), H_r(Y))$$

- Which one is better?
  Neither
- Does Theorem 2 hold under Condition (A)?
  No (counterexample)
- The EMCEL scheme satisfies Condition (D), hence Condition (B)
  $\Longrightarrow$ **every** general diffusion can be approximated by EMCEL **together with its exit times**

A specific open question

- Chigansky–Klebaner (2012) consider the CEV diffusion

$$dY_t = Y_t^p \, dW_t$$

with state space $\mathcal{I} = [0, \infty)$ and parameter $p \in [1/2, 1)$, its strong Euler approximations $X^{stEu,h}$

- They prove

$$H_{h^\beta}(X^{stEu,h}) \xrightarrow[h \searrow 0]{w} H_0(Y)$$

for any fixed $\beta \in \left(0, \frac{1/2}{1-p}\right)\ldots$

- ...and announce an open question of whether it is true that

$$H_0(X^{stEu,h}) \xrightarrow[h \searrow 0]{w} H_0(Y)$$

Answer for the weak Euler scheme

- Weak Euler approximations $X^{wEu,h}$ of the CEV diffusion do approximate its exit time. Moreover,

$$(X^{wEu,h}, H_0(X^{wEu,h})) \xrightarrow[h \searrow 0]{w} (Y, H_0(Y)).$$

  Reason: the weak Euler scheme for the CEV diffusion satisfies Condition (B)

- Notice: the weak Euler scheme for the CEV diffusion does not satisfy Condition (D)

Intermediate summary

- The EMCEL scheme satisfies Condition (D), hence Condition (B) $\Longrightarrow$ **every** general diffusion can be approximated by EMCEL **together with its exit times**
- In case when the equation determining EMCEL cannot be solved in closed form, Condition (B) dictates the required precision for the numerical solution in order to retain the convergence of processes & exit times
- Alternatively, Condition (B) can be used to verify convergence of processes & exit times for other schemes, e.g., for the weak Euler one

Rate of convergence

Preprint on arXiv:

- Ankirchner, Kruse, Urusov. “Wasserstein convergence rates”, 2019

Condition on the general diffusion

Condition (C): there exist constants $k_1 \in (0, \infty)$, $k_2 \in \{0, 1\}$ such that

$$m(dx) \geq \frac{2}{k_1(1 + k_2 x^2)}\, dx$$

- In the SDE case

$$dY_t = \eta(Y_t)dW_t, \quad m(dx) = \frac{2}{\eta^2(x)}\, dx,$$

  Condition (C) is just linear growth condition for η

- But it is not only for the SDE case: e.g., Condition (C) allows "stickiness"

- Contrary to Conditions (A), (B) and (D), which are requirements on the approximation scheme (automatically satisfied for EMCEL), Condition (C) is a requirement on the general diffusion

Condition on the scheme (idea)

Condition (Aλ): there exists $\lambda \in (0, \infty)$ such that the scale factors $(a_h)_{h \in (0, \bar{h})}$ satisfy

$$\sup_{y \in \mathcal{I}} \left| \frac{1}{2} \int_{(y - a_h(y), y + a_h(y))} (a_h(y) - |u - y|)\, m(du) - h \right| = O(h^{1+\lambda}), \quad h \to 0$$

- Condition (Aλ) implies Condition (A)
- The EMCEL scheme satisfies Condition (Aλ) for all $\lambda \in (0, \infty)$

Wasserstein distances

Let $p \in [1, \infty)$

- Spaces: let $\mathcal{M}_p(\mathbb{R})$ (resp. $\mathcal{M}_p(C([0,T],\mathbb{R}))$) be the set of probability measures on $\mathbb{R}$ (resp. on $C([0,T],\mathbb{R})$) with finite p-th moment (e.g., $\int \sup_{t\in[0,T]} |x(t)|^p \, \mu(dx) < \infty$ for $C([0,T],\mathbb{R})$)
- p-th Wasserstein distance between $\mu, \nu \in \mathcal{M}_p(\mathbb{R})$:

$$\mathcal{W}_p(\mu,\nu) = \inf \|\xi - \zeta\|_{L^p},$$

  where the infimum is taken over random vectors (ξ, ζ) with $\xi \sim \mu$ and $\zeta \sim \nu$
- p-th Wasserstein distance between $\mu, \nu \in \mathcal{M}_p(C([0,T],\mathbb{R}))$:

$$\mathcal{W}_p(\mu,\nu) = \inf \left\| \sup_{t\in[0,T]} |\xi_t - \zeta_t| \right\|_{L^p}$$

Rates for time marginals and on the path space

Theorem (Rate for time marginals)

Under Condition (C) and Condition (Aλ) there exists $C(p,T) \in [0,\infty)$ such that for all $y \in \mathcal{I}^\circ$ and $h \in (0,\overline{h})$ it holds

$$\mathcal{W}_p(P \circ (X_T^{h,y})^{-1}, P_y \circ (Y_T)^{-1}) \le C(p,T)(1 + k_2|y|)h^{\min\{\frac{1}{4},\frac{\lambda}{2}\}}$$

Theorem (Rate on the path space)

Let $\varepsilon > 0$. Under Condition (C) and Condition (Aλ) there exists $C(p,\varepsilon,T) \in [0,\infty)$ such that for all $y \in \mathcal{I}^\circ$ and $h \in (0,\overline{h})$ it holds

$$\mathcal{W}_p(P \circ (X^{h,y})^{-1}, P_y \circ (Y)^{-1}) \le C(p,\varepsilon,T)(1 + k_2|y|)h^{\min\{\frac{1}{4},\frac{\lambda}{2}\}-\varepsilon}$$

Role of Condition (C)

- Condition (C) guarantees that Law(Y) belongs to $\mathcal{M}_p(C([0,T],\mathbb{R}))$ for all $p \in [1,\infty)$
- We have counterexamples when Condition (C) is violated

Moreover, as we have rates (hence convergence) in **every** p-th Wasserstein distance, we get

Corollary

Under Condition (C) and Condition (Aλ) we have

$$\mathsf{E}\left[F(X_t^{h,y}; t \in [0,T])\right] \to \mathsf{E}_y\left[F(Y_t; t \in [0,T])\right], \quad h \to 0, \qquad (*)$$

for continuous path functionals F ***with polynomial growth***

- Essentially, Condition (C) is the price for a stronger mode of convergence (recall: under Condition (A) we have weak convergence, which is $(*)$ for **bounded** continuous path functionals)

Intermediate summary

- The EMCEL scheme satisfies Condition (Aλ) for all $\lambda \in (0, \infty)$ $\Longrightarrow$ **every** general diffusion **satisfying Condition (C)** can be approximated by EMCEL with rate $1/4$
- Technically, Condition (C) ensures that Law(Y) belongs to the right space, and it can be viewed as a price for a stronger mode of convergence (than the weak convergence on the path space)
- In case when the equation determining EMCEL cannot be solved in closed form, Condition (Aλ) dictates the required precision for the numerical solution in order to have a rate:

  More precisely, the precision of $O(h^{3/2})$ still provides rate $1/4$, while a further reduction of the precision in solving the equation determining EMCEL leads to smaller rates

How good is the rate of $1/4$?

- SDE case, $\eta \in C^4$ with bounded derivatives $\Longrightarrow$ the Euler scheme has rate 1

  [Kloeden–Platen (1992)]

- SDE case, η Hölder with exponent $\gamma \Longrightarrow$ the Euler scheme has rate $\gamma/2$

  [Konakov–Menozzi (2017), Frikha (2018)]

$\longrightarrow$ Rate $1/4$ is good for irregular cases

Rate for approximating expected values

Corollary

Suppose Condition (C) and Condition (Aλ). Let $\varepsilon > 0$, $\alpha, L \geq 0$ *and* $F\colon C([0,T],\mathbb{R}) \to \mathbb{R}$ *satisfy for all* $x_1, x_2 \in C([0,T],\mathbb{R})$

$$|F(x_1) - F(x_2)| \leq L\left\{1 + (\|x_1\|_C \vee \|x_2\|_C)^\alpha\right\}\|x_1 - x_2\|_C.$$

Then there exist $C(\alpha,\varepsilon,T) \in [0,\infty)$ *such that for all* $h \in (0,\overline{h})$ *and* $y \in \mathcal{I}^\circ$ *it holds*

$$\left|E\left[F(X_t^{h,y};\, t\in[0,T])\right] - E_y\left[F(Y_t;\, t\in[0,T])\right]\right| \leq LC(\alpha,\varepsilon,T)(1+|y|^\alpha)(1+k_2|y|)h^{\min\{\frac{1}{4},\frac{\lambda}{2}\}-\varepsilon}$$

Summary

What was presented

On stochastic side:

- FLT for weak convergence on the path space
- FLT for processes & exit times
- Rate of convergence for time marginals and on the path space in every p-th Wasserstein distance

On numerical side:

- Numerical scheme EMCEL that works for **all** general diffusions (for Wasserstein need Condition (C))
- All results include perturbation analyses

Thank you!

Preprints on arXiv:

- Ankirchner, Kruse, Urusov. “Functional limit theorem”, 2019
- Kruse, Urusov. “Exit times”, 2019
- Ankirchner, Kruse, Urusov. “Wasserstein convergence rates”, 2019