Ordre convexe fonctionnel pour les processus stochastiques : une approche constructive (et simulable)

Gilles Pagès

(including joint works with B. Jourdain, Y. Liu & C. Yeo)

LPSM-Sorbonne-Université

Cours Bachelier

26 avril 2024

Outline

Definitions

Definition (Convex orderings)

Let $U,\ V \in L^1_{\mathbb{R}^d}(\mathbb{P})$ be two $\mathbb{R}^d$-valued random vectors with distributions μ and ν.
(a) *Convex ordering*. We say that U is dominated for the convex ordering by V, denoted

$$U \preceq_{cvx} V$$

if, for every convex function $f : \mathbb{R}^d \to \mathbb{R}$,

$$\mathbb{E}\, f(U) \leq \mathbb{E}\, f(V) \in (-\infty, +\infty] \tag{1}$$

or, equivalently, that μ is dominated by ν for the convex ordering if, for every convex function $f : \mathbb{R}^d \to \mathbb{R}$, $\int_{\mathbb{R}^d} f\, d\mu \leq \int_{\mathbb{R}^d} f\, d\nu$.

(b) *Monotone convex ordering ($d = 1$)*. When (1) only holds for non-decreasing/non-increasing convex functions f, the convex ordering is called increasing/decreasing convex order respectively denoted

$$U \preceq_{icv} V \quad \text{and} \quad U \preceq_{dcv} V.$$

Consistency

- For every $x \in \mathbb{R}^d$, by convexity of $f : \mathbb{R}^d \to \mathbb{R}$,

$$f(x) \geq f(0) + \langle \nabla_s f(0) \,|\, x \rangle.$$

where $\nabla_s f(0)$ denotes a subgradient of f at 0.

- Hence

$$f^-(x) \leq \Big(f(0) + \langle \nabla_s f(0) \,|\, x \rangle \Big)^- \leq |f(0)| + |\nabla_s f(0)||x|$$

so that

$$\mathbb{E}\, f^-(U) \leq |f(0)| + |\nabla_s f(0)| \mathbb{E}\,|U| < +\infty.$$

Consequently

$$\mathbb{E}\, f(U) = \underbrace{\mathbb{E}\, f^+(U)}_{\in[0,+\infty]} - \underbrace{\mathbb{E}\, f^-(U)}_{\in[0,+\infty)} \in (-\infty, +\infty] \text{ is well-defined.}$$

First properties (of $\preceq_{cvx}$)

- **P1**. As $f(x_1,\ldots,x_d)) = \pm x_i$, $i = 1:d$, are all convex, $U \preceq_{cvx} V$ implies
$$\mathbb{E}\,U = \mathbb{E}\,V.$$
- **P2**. If, U, $V \in L^2(\mathbb{P})$, $U \preceq_{cvx} V$, then we have with $f(x) = |x|^2$
$$\mathrm{Var}(U) \le \mathrm{Var}(V)$$
[where $\mathrm{Var}(U) = \mathbb{E}\,|U|^2 - |\mathbb{E}\,U|^2$].
- **P3**. If $U \preceq_{icv} V$, then $\mathbb{E}\,U \le \mathbb{E}\,V$.
- **P4**.
$$U \preceq_{dcv} V \Longleftrightarrow -U \preceq_{icv} -V$$
since $f(x) = f\big(-(-x)\big)$ and $f(-\cdot)$ is convex with opposite monotony.

Convex ordering is a kind of generalization of the measure of risk

through the variance.

Examples I

- If $U = \mathbb{E}\left(V \mid U\right)$ then, for every convex function such that $f(V) \in L^1(\mathbb{P})$,

$$\mathbb{E}\, f(U) = \mathbb{E}\, f\big(\mathbb{E}\,(V \mid U)\big) \leq \mathbb{E}\left[\mathbb{E}\,(f(V) \mid U)\right] = \mathbb{E}\, f(V).$$

owing to Jensen's inequality. Obvious if $\mathbb{E}\, f(V) = +\infty$.

- If $U \perp\!\!\!\perp W$, $W \in L^1(\mathbb{P})$, $\mathbb{E}\, W = 0$, then $U \preceq_{cvx} U + W$. $[\mu \preceq_{cvx} \mu * \nu_0]$
- $\delta_{\mathbb{E} V} \preceq_{cvx} V$. $[\delta_{\int \xi\, \nu(d\xi)} \preceq_{cvx} \nu]$
- Gaussian distributions (centered): Let $Z \sim \mathcal{N}(0, I_q)$ on $\mathbb{R}^q$ and let A, $B \in \mathbb{M}_{d,q}$ be $d \times q$ matrices

$$AA^* \leq BB^* \;\text{ in }\; \mathcal{S}^+(d, \mathbb{R}) \Longleftrightarrow AZ \preceq_{cvx} BZ$$

or equivalently $\mathcal{N}(0, AA^*) \preceq_{cvx} \mathcal{N}(0, BB^*)$.

In particular if $d = q = 1$, $|\sigma| \leq |\vartheta| \Longleftrightarrow \mathcal{N}(0, \sigma^2) \preceq_{cvx} \mathcal{N}(0, \vartheta^2)$.

- Proof. Let $Z_1, Z_2 \sim \mathcal{N}(0; I_q)$ be independent. Set

$$U = AZ_1, \quad V = U + (BB^* - AA^*)^{1/2} Z_2.$$

Then $U = \mathbb{E}\left(V \mid U\right)$ and $V \sim \mathcal{N}\Big(0, AA^* + \big((BB^* - AA^*)^{1/2}\big)^2\Big) = \mathcal{N}(0, BB^*)$.

- 1D-proof: $\varphi : \mathbb{R} \to \mathbb{R}$ convex and $Z \in L^1$, $Z \stackrel{d}{=} -Z$. Then, by Jensen's $\leq$,

$$u \mapsto \mathbb{E}\,\varphi(uZ) \text{ is even, convex and attains its minimum } \varphi(0) \text{ at } u = 0.$$

Hence $u \mapsto \mathbb{E}\,\varphi(uZ)$ is non-decreasing on $\mathbb{R}_+$ and non-increasing on $\mathbb{R}_-$.

- Generalization: radial distributions: Let $Z : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{R}^q$ having a radial distribution in the sense

$$\forall\, O \in \mathcal{O}(q), \quad OZ \sim Z.$$

Let A, $B \in \mathbb{M}_{d,q}$. Then

$$AA^* \leq BB^* \ \text{ in } \ \mathcal{S}^+(d, \mathbb{R}) \Longrightarrow AZ \preceq_{cvx} BZ$$

We skip the proof (exercise with solution in ([1])).

[1] B. Jourdain, G. Pagès, Convex order, quantization and monotone approximations of ARCH models, *Journal of Theoretical Probability*, 35, (4), 2480–2517,2022

- If $U \preceq_{cvx} V$ and $U' \preceq_{cvx} V'$, $U \perp\!\!\!\perp U'$, $V \perp\!\!\!\perp V'$ then

$$U + U' \preceq_{cvx} V + V'.$$

[$\mu \preceq_{cvx} \nu$ and $\mu' \preceq_{cvx} \nu' \Rightarrow \mu * \mu' \preceq_{cvx} \nu * \nu'$]. By Fubini's Theorem

$$\begin{aligned}\mathbb{E}\, f(U+U') &= \int_{\mathbb{R}^d} \mathbb{E}\, f(u+U')\mathbb{P}_U(du) \le \int_{\mathbb{R}^d} \mathbb{E}\, f(u+V')\mathbb{P}_U(du)\\ &\le \int_{\mathbb{R}^d} \mathbb{E}\, f(u+V')\mathbb{P}_{U'}(du) = \mathbb{E}\, f(U'+V').\end{aligned}$$

- If $(U_n)_{n\ge1}$ i.i.d.$\sim U$ and $(V_n)_{n\ge1}$ i.i.d.$\sim V$, centered, $\perp\!\!\!\perp N, M$, $N \le M$, having values in $\mathbb{N}_0$, integrable

$$\sum_{k=1}^{N} U_k \preceq_{cvx} \sum_{k=1}^{N} V_k \preceq_{cvx} \sum_{k=1}^{M} V_k.$$

Obvious by induction when N and M are integers, etc.

Example II: martingales, peacocks

- If $(X_t)_{t\geq 0}$ is a martingale, then

$$t \longmapsto X_t \text{ is non-decreasing for the convex ordering}$$

i.e. $0 \leq s \leq t \Rightarrow X_s \preceq_{cvx} X_t$ since

$$\forall\, 0 \leq s \leq t, \quad X_s = \mathbb{E}\left(X_t \,|\, X_s\right).$$

- More generally, a process such that

$$t \longmapsto X_t \text{ is non-decreasing for the convex ordering}$$

is called p.c.o.c (for “Processus Croissant pour l’Ordre Convexe” in French) or even ”peacock”...).

- Thus, any martingale is a peacock !
- More generally, if $X_t \sim M_t$, $t \geq 0$, where $(M_t)_{t\geq 0}$ is a martingale, then $(X_t)_{t\geq 0}$ is a peacock

About converses of "$U = \mathbb{E}(V \mid U) \Rightarrow U \preceq_{cvx} V$" and "1-martingale $\Rightarrow$ p.c.o.c."

- Strassen's Theorem (1965): $\mu \preceq_{cvx} \nu \Longleftrightarrow \exists$ transition $P(x, dy)$ s.t.

$$\nu = \mu P \quad \text{and} \quad \forall\, x \in \mathbb{R}^d, \quad \int y\, P(x, dy) = x.$$

- Kellerer's Theorem (1972):

$$X \text{ is a p.c.o.c },\Longleftrightarrow \exists\,(M_t)_{t\geq 0} \text{ such that } X_t \overset{d}{=} M_t,\; t \geq 0,$$

  (X is sometimes called a "1-martingale").
- Both proofs are unfortunately non-constructive.
- In Hirsch, Roynette, Profeta & Yor's monography ([2]), many (many...) explicit "representations" of p.c.o.c. by true martingales. Also, investigations on 2-martingales, n-martingales...

[2] *Peacocks and Associated Martingales, with Explicit Constructions*, Springer, 2011.

A revival motivated by Finance...

- A starter! t being fixed, $\sigma \mapsto e^{\sigma W_t - \frac{\sigma^2 t}{2}}$ is a p.c.o.c. since

$$\forall \sigma > 0, \quad e^{\sigma W_t - \frac{\sigma^2 t}{2}} \stackrel{d}{=} e^{W_{\sigma^2 t} - \frac{\sigma^2 t}{2}} \; (\to \text{martingale w.r.t. } \sigma).$$

- Application to Black-Scholes model $S_t^\sigma = s_0 e^{\sigma W_t - \frac{\sigma^2 t}{2}}$. For every convex payoff function $f : \mathbb{R}_+ \to \mathbb{R}_+$

$$0 < \sigma \leq \sigma' \Longrightarrow \mathbb{E}\, f(S_t^\sigma) \leq \mathbb{E}\, f(S_t^{\sigma'}).$$

- Vanilla options: *Call* and *Put* options: $f(S_T) = (S_T - K)^+$, $f(S_T) = (K - S_T)^+$, etc.
- In fact, $U \preceq_{icvx} V$ iff $\forall K \in \mathbb{R}$, $\mathbb{E}\,(U-K)^+ \leq \mathbb{E}\,(V-K)^+$
- and $U \preceq_{cvx} V$ iff $\mathbb{E}\, U = \mathbb{E}\, V$ and $\forall K \in \mathbb{R}$, $\mathbb{E}\,(U-K)^+ \leq \mathbb{E}\,(V-K)^+$.

Path-dependent payoffs

- E.g. what about path-dependent options like Asian payoffs. Let $f : \mathbb{R}_+ \to \mathbb{R}_+$ convex

$$\sigma \longmapsto \mathrm{Premium}(\sigma) = \mathbb{E}\Big[f\Big(\frac{1}{T}\int_0^T \underbrace{s_0 e^{\sigma W_t - \frac{\sigma^2 t}{2}}}_{=S_t^\sigma} dt\Big)\Big] \,?$$

- P. Carr et al. (2008):
  Non-decreasing in σ when $f(x) = (x-K)^+$ (Asian Call).
- Baker-Yor (2010):

$$\sigma \mapsto \frac{1}{T}\int_0^T s_0 e^{\sigma W_t - \frac{\sigma^2 t}{2}} dt \stackrel{(t=uT)}{=} \int_0^1 s_0 e^{W_{u\sigma^2 T} - \frac{u\sigma^2 T}{2}} du \stackrel{d}{=} \int_0^1 s_0 e^{W_{u,\sigma^2 T} - \frac{u\sigma^2 T}{2}} du$$

  where $(W_{u,t})_{u,t\geq 0}$ is a standard Brownian sheet. Hence a p.c.o.c. since

$$t \mapsto \int_0^1 s_0 e^{W_{u,t} - \frac{ut}{2}} du \text{ is an } (\mathcal{F}_{1,t})_{t\geq 0}\text{-martingale.}$$

- Yields bounds on the option prices of vanilla options:

$$\boxed{\sigma_{\min} \leq \sigma \leq \sigma_{\max} \Longrightarrow \text{etc.}}$$

▷ This suggests many other (new or not so new) questions !

- Non-decreasing convex ordering: $\exists$ drift b! [see [Hajek, 1985] ([3]).
- Functional convex order I: Switch from BS to local volatility models *i.e.* from scalar (or vector) parameter to a functional parameter.

$$\sigma \rightsquigarrow \sigma(x)$$

[see e.g. El Karoui-Jeanblanc-Schreve, 1998] i.e.

$$dX_t = \sigma(X_t)dW_t,\ X_0 \perp\!\!\!\perp W \ \text{ vs } \ dY_t = \theta(Y_t)dW_t,\ Y_0 \perp\!\!\!\perp W,\ X_0 \preceq_{cvx} Y_0, \ldots ?$$

- m-marginal path-dependent convex order: $f(X_t) \rightsquigarrow F(X_{t_1}, \ldots, X_{t_m})$ [see Brown, Rogers, Hobson 2001, Rüschendorf et al., 2008].
- Functional convex order II : from $\mathbb{E}\, f(X_T^{(\sigma)})$ to $\mathbb{E}\, F(X^{(\sigma)})$ path-dependent convex order, [see P.2016].
- Bermuda and American options [see Pham 2005, Rüschendorf 2008, P. 2016].
- Jumpy risky asset dynamics for (X_t^{σ})? [see Rüschendorf-Bergenthum, 2007, P. 2016]).
- P.c.o.c. trough Martingale Optimal Transport. [see Beïgelbock, Henry-Labordère et al, 2013, Tan et al. 2015, Jourdain-P. 2020].

[3] Hajek, B., Mean stochastic comparison of diffusions. Z. Wahrsch. Verw. Gebiete 68 (1985), no. 3, 315–329.

More questions about convexity

- A side (?) question of interest : propagation of convexity in the sense

$$f:\mathbb{R}\to\mathbb{R} \text{ convex } \Longrightarrow x\longmapsto \mathbb{E}\, f(X_T^x) \text{ convex ?}$$

e.g. in a1D- local volatility model like

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

- More generally, when do we have such propagation of convexity if

$$X_t^x = x + \int_0^t \alpha(X_s^x + \beta) ds + \int_0^t \sigma(s, X_s^x) dW_s \quad ?$$

- Extensions to convex functionals $F:\mathcal{C}([0,T],\mathbb{R})\to\mathbb{R}$ and to higher dimensional processes ($d\geq 2$) ?
- Similar questions for monotonic convexity with a more general drift

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

Direct approach: first reduction

- Assume $\sigma(t,y)$ Lipschitz in y uniformly in $t\in[0,T]$ and $\sigma(\cdot,0)$ bounded.
- Let $f:\mathbb{R}\to\mathbb{R}$ be convex

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

- Setting

$$\widetilde{X}_t^x = e^{\alpha t}X_t - \beta(1-e^{\alpha t})$$

and

$$\widetilde{\sigma}(t,y) = e^{\alpha t}\sigma\big(t, e^{-\alpha t}y - \beta(1-e^{-\alpha t})\big)$$

yields

$$\widetilde{X}^x = x + \int_0^t \widetilde{\sigma}(s,\widetilde{X}_s^x)dW_s$$

where $\widetilde{\sigma}(t,y)$ Lipschitz in y uniformly in $t\in[0,T]$.
- Hence, we may assume w.l.g. $\alpha=\beta=0$.

Aims and methods

1. Unify and generalize existing results with of focus on both functional aspects of functional convex ordering.
   - with a focus on both functional aspects of functional convex ordering.
   - As a by-product establish the convexity of $x \mapsto \mathbb{E}\, f(X_T^x)$ and/or $x \mapsto \mathbb{E}\, F(x^x)$.
2. Constraint: provide a constructive method of proof.
   - based on time discretization of continuous time martingale dynamics (risky assets in Finance) .
   - using numerical schemes that preserve the functional convex order satisfied by the process under consideration...
   - to avoid arbitrages.
3. Apply the paradigm to various frameworks:
   - American style options,
   - jump diffusions,
   - stochastic integrals,
   - McKean-Vlasov diffusions,
   - Volterra equations,
   - etc?

Example III: risk measure

- Let $X \in L^1\mathbb{P}$ be representative of a loss (with no atom for convenience) with c.d.f F_X.
- Let $\alpha \in (0,1]$, $\alpha \simeq 1$ be a risk level. Then

$$\mathrm{VaR}_\alpha(X) := (F_X)^{-1}(\alpha) \quad \text{and} \quad \mathrm{CVaR}_\alpha(X) := \mathbb{E}\big(X \,|\, X \geq \mathrm{Var}_\alpha(X)\big)$$

- Rockafeller-Uryasev's representation of these two risk measures

$$L_{\alpha,X}(\xi) = \xi + \frac{1}{1-\alpha}\mathbb{E}\,(X-\xi)^+$$

satisfies

$$\mathrm{Var}_\alpha(X) = \mathrm{argmin}_{\mathbb{R}} L_{\alpha,X} \quad \text{and} \quad \mathrm{CVaR}_\alpha(X) = \min_{\mathbb{R}} L_{\alpha,X}.$$

- As a consequence

$$X \preceq_{icv} Y \Longrightarrow L_{\alpha,X} \leq L_{\alpha,Y}$$

so that

$$\mathrm{CVaR}_\alpha(X) \leq \mathrm{CVaR}_\alpha(Y).$$

- WARNING! Not true for the value-at-risk.

Characterization of convex ordering

Proposition

(a) Let $U,\ V \in L^1_{\mathbb{R}^d}(\mathbb{P})$. There is equivalence between

$$U \preceq_{cvx} V$$

and

$$\forall f : \mathbb{R}^d \to \mathbb{R} \text{ convex and Lipschitz continuous } \mathbb{E}\, f(U) \leq \mathbb{E}\, f(V).$$

(b) Similar equivalence for $\preceq_{icv}$ and $\preceq_{dcv}$ (when $d = q = 1$).

The proof relies on the following lemma based on inf-convolution.

Lemma (Approximate convex functions from below)

Any convex function $f : \mathbb{R}^d \to \mathbb{R}$ satisfies

$$f = \lim_n {}^{\uparrow} f_n, \quad f_n \text{ convex and Lipschitz continuous}, \quad n \geq 1.$$

The functions f_n have the same monotonicity as f, if any.

Proof (lemma)

- We introduce the functions f_n defined through inf-convolution on $\mathbb{R}^d$ by

$$f_n(x) := \inf_{y\in\mathbb{R}^d} \big(f(y) + n|x-y|\big),\; n \geq 1.$$

One has by construction

$$\forall\, n \geq 1, \quad f_n \leq f_{n+1} \leq f.$$

- $f_n \uparrow f$ in a stationary way: let us denote by $\nabla_s f(x)$ any subgradient of f at x.

$$\begin{aligned}\forall\, y\in\mathbb{R}^d,\; f(y) + n|y-x| &\geq f(x) + \langle \nabla_s f(x)\,|\,y-x\rangle + n|y-x| \text{ by convexity of } f\\ &\geq f(x) + (n - |\nabla_s f(x)|)|y-x|\\ &\geq f(x)\end{aligned}$$

Hence, $\forall\, n \geq |\nabla_s f(x)|, \quad f_n(x) \geq f(x)$ so that $f_n(x) = f(x)$.

- f_n is convex since, for $x, x'\in\mathbb{R}^d$, $\lambda\in[0,1]$,

$$\begin{aligned} f_n(\lambda x + (1-\lambda)x') &= \inf_{y,y'} f(\lambda y + (1-\lambda)y')) + n|\lambda(x-y) + (1-\lambda)(x'-y')|\\ &\leq \lambda \inf_y \big(f(y) + n|x-y|\big) + (1-\lambda)\inf_{y'}\big(f(y') + n|x'-y'|\big)\\ &= \lambda f_n(x) + (1-\lambda) f_n(x').\end{aligned}$$

Proof ($\Leftarrow$ of proposition)

- f_n are n-Lipschitz continuous since

$$|f_n(x) - f_n(x')| \le \sup_{y\in\mathbb{R}^d} \big| n|x-y| - n|x'-y| \big| \le n|x-x'|.$$

- $f_n(x) = \inf_y \big(f(x+y) + n|y|\big)$ has the same monotonicity as f ... if any. $\square$

Proof of the proposition.

- Assume f convex, then for every $n \ge 1$, $\mathbb{E}\, f_n(U) \le \mathbb{E}\, f_n(V)$.
- The functions f_n^-, $n \ge |\nabla_s f(0)|$, are dominated since

$$\begin{aligned}\forall\, x,\, y \in \mathbb{R}^d,\ f_n(x) &\ge f(0) + \langle \nabla_s f(0)\,|\, y\rangle + n|y-x|.\\ &\ge f(0) + |y|(n - |\nabla_s f(0)|) - n|x| \ge f(0) - n|x|.\end{aligned}$$

- As $U,\ V \in L^1(\mathbb{P})$, one has by the monotone convergence theorem

$$-\infty < \mathbb{E}\, f(U) \le \mathbb{E}\, f(V) \le +\infty. \qquad \square$$

Functional convex ordering: Definition

Assume $\mathcal{C}_T = \mathcal{C}([0,T],\mathbb{R}^d)$ is equipped with sup-norm $\|f\|_{\sup} = \sup_{u\in[0,T]} |f(u)|$.

Definition

Let $X, Y : (\Omega, \mathcal{F}, \mathbb{P}) \to \mathcal{C}([0,T],\mathbb{R}^d)$ be two integrable continuous processes such that $\mathbb{E}[\|X\|_{\sup} + \|Y\|_{\sup}] < +\infty$.

(a) *Convex ordering.* We say that X is dominated by Y for the *convex ordering* – denoted by $X \preceq_{cvx} Y$ – if, for every l.s.c. (for the $\|\cdot\|_{\sup}$-norm topology) convex functional $F : \mathcal{C}([0,T],\mathbb{R}^d) \to \mathbb{R}$,

$$\mathbb{E}\, F(X) \leq \mathbb{E}\, F(Y). \tag{2}$$

(b) *Monotone convex ordering $(d=1)$.* We say that X is dominated by Y for the *increasing/decreasing convex ordering* if (2) holds for every non-increasing/non-decreasing for the pointwise partial order on $\mathcal{C}$ l.s.c. convex functional $F : \mathcal{C}([0,T],\mathbb{R}) \to \mathbb{R}$. These orderings are denoted by

$$X \preceq_{icv} Y \quad \text{and} \quad X \preceq_{dcv} Y \quad \text{respectively.}$$

Characterization of functional convex ordering

- Do we have the same characterization by Lipschitz functionals ? Yesss!

Proposition

Let X, Y be two $\mathcal{C}([0,T],\mathbb{R}^d)$-valued r.v. (i.e. pathwise continuous stochastic processes) such that $\mathbb{E}[\|X\|_{\sup} + \|Y\|_{\sup}] < +\infty$.

(a) *Convex order*. Both statements are equivalent:

$$X \preceq_{cvx} Y$$

and

$$\forall\, F \in \mathcal{C}([0,T],\mathbb{R}^d) \to \mathbb{R},\ \|\cdot\|_\infty\text{-Lipschitz continuous},\ \mathbb{E}\,F(X) \le \mathbb{E}\,F(Y).$$

(b) *Pointwise monotonic convex ordering $(d=1)$*. Similar equivalence for $X \preceq_{icv} Y$ and $X \preceq_{dcv} Y$ with respect to pointwise non-decreasing (resp. non-increasing) Lipschitz convex functionals $F : \mathcal{C}([0,T],\mathbb{R}^d) \to \mathbb{R}$.

- The key is the following miracle-lemma!

Miracle lemma

Lemma (Quasi-subgradient)

([a]) Let $(E, \|\cdot\|)$ be a normed vector space and let $F : E \to \mathbb{R}$ be an l.s.c. convex functional (for the norm topology).

For every $x \in E$ and every $a \in (-\infty, F(x))$; there exists $G = G_{x,a} \in E'$ and $g = g_{x,a} \in \mathbb{R}$ such that

$$\begin{array}{ll} (i) & \forall\, u \in E, \quad G(u) + g \le F(u), \\ (ii) & G(x) + g = a. \end{array}$$

[a] See Lemma 7.5 in Aliprantis, Charalambos D. and Border, Kim C., *Infinite dimensional Analysis*, Springer, 2006.

- The linear forms $G_{x,a}$, $-\infty < a < F(x)$ play the role of the sub gradient and the characterization in $\mathbb{R}^d$ can be extended to this framework with $E = \mathcal{C}([0, T], \mathbb{R}^d)$.
- One shows likewise that $\mathbb{E}\, F(X) \in (-\infty, +\infty]$ and the characterization by Lipschitz continuous functionals.

Paradigm of convex ordering by Wasserstein approximation

- Let $(E, |\cdot|_E)$ be a Banach space and

$$\mathcal{P}_1(E) = \Big\{\mu \text{ distribution on } \big(E, \mathcal{B}or(E)\big) : \int_E |\xi|_E \mu(d\xi) < +\infty\Big\}$$

be the convex set of integrable probability measures on $(E, \mathcal{B}or(E))$ equipped with the (metric) topology of $\mathcal{W}_1$ the Wasserstein/Monge-Kantorovich distance.

$$\mathcal{W}_1(\mu,\nu) = \inf\Big\{\int |x-y| m(dx,dy),\ m(dx,E)=\mu,\ m(E,dy)=\nu\Big\} = \sup\Big\{\int f d\mu - \int f d\nu, [f]_{\rm Lip} \le 1\Big\}.$$

- Let X and Y be two E-valued random variables and let $(X_n)_{n\ge1}$ and $(Y_n)_{n\ge1}$ two sequences of E-valued random variables such that

$$(i) \quad \forall\, n \ge 1, \quad X_n \preceq_{cvx} Y_n$$

$$(ii) \quad \mathcal{W}_1([X_n],[X]) + \mathcal{W}_1([Y_n],[Y]) \to 0 \text{ as } n \to +\infty$$

where $[X] \in \mathcal{P}_1(E)$ denotes the distribution of X. Then

$$X \preceq_{cvx} Y.$$

Proof of the paradigm

- Let $F : E \to \mathbb{R}$ be a Lipschitz continuous function. Assumption (i) implies that
$$\mathbb{E}\, F(X_n) \le \mathbb{E}\, F(Y_n), \quad n \ge 1.$$
- Then, by (ii) and the Monge-Kantorovich characterization of $\mathcal{W}_1$-distance
$$\big|\mathbb{E}\, F(X_n) - \mathbb{E}\, F(X)\big| \le [F]_{\rm Lip}\mathcal{W}_1\big([X_n],[X]\big) \to 0 \text{ as } n \to +\infty,$$
- Idem for Y_n and Y.
- Letting $n \to +\infty$ in the first inequality yields the conclusion. □

▷ Application to $E = \mathcal{C}([0,T], \mathbb{R}^d), \|\cdot\|_{\sup})$.

▷ Adaptation to partially-ordered Banach space is straightforward.

▷ Other extensions e.g. to metric vector spaces (think to Skorokhod topology on $\mathbb{D}([0,T], \mathbb{R}^d)$.)

Martingale (and scaled) Brownian diffusions

- If we want to compare on (l.s.c.) convex functionals $F : \mathcal{C}([0,T],\mathbb{R}^d) \to \mathbb{R}$,

$$\mathbb{E}\, F(X) \quad ? \quad \mathbb{E}\, F(Y)$$

where

$$dX_t = \sigma(t, X_t)dW_t,\; X_0 \perp\!\!\!\perp W \;\text{ versus }\; dY_t = \theta(t, Y_t)dW_t,\; Y_0 \perp\!\!\!\perp W,\; X_0 \preceq_{cvx} Y_0?$$

in a higher dimensional setting:

– W q-dimensional B.M.,

– $\sigma(t,\cdot) : \mathbb{R}^d \to \mathbb{M}_{d,q}(\mathbb{R})$

we need:

- a pre-order on matrices,
- the resulting notion of convexity for matrix-valued vector fields.

Martingale (and scaled) Brownian diffusions

- Pre-order $\preceq$ on $\mathbb{M}_{d,q}(\mathbb{R})$: let $A,\ B\in\mathbb{M}_{d,q}(\mathbb{R})$.

$$A\preceq B \quad \text{if} \quad BB^*-AA^*\in\mathcal{S}^+(d,\mathbb{R}).$$

[If $d=q=1$, $a\preceq b$ iff $|a|\leq|b|$].

- $\preceq$-Convexity: $\sigma:\mathbb{R}^d\to\mathbb{M}_{d,q}$ is $\preceq$-convex if

$\forall\, x,\ y\in\mathbb{R}^d$, $\lambda\in[0,1]$, there exists $O_{\lambda,x},\ O_{\lambda,y}\in\mathcal{O}(q,\mathbb{R})$ such that

$$\sigma\big(\lambda x+(1-\lambda)y\big)\preceq\lambda\sigma(x)O_{\lambda,x}+(1-\lambda)\sigma(y)O_{\lambda,y}$$

i.e.

$$\sigma\sigma^*\big(\lambda x+(1-\lambda)y\big)\leq\big(\lambda\sigma(x)O_{\lambda,x}+(1-\lambda)\sigma(y)O_{\lambda,y}\big)\big(\lambda\sigma(x)O_{\lambda,x}+(1-\lambda)\sigma(y)O_{\lambda,y}\big)^*$$

- $d=q=1$ with $O_{\lambda,x}=\operatorname{sign}\big(\sigma(x)\big)$ this simply reads

$$|\sigma|\,\mathit{convex}.$$

- $\Longrightarrow$ WARNING! Then, if $d=q=1$: σ $\preceq$-convex means $|\sigma|$ convex !!

Examples

- Let $\lambda_k : \mathbb{R} \to \mathbb{R}$, $k = 1 : q$ be Lipschitz functions such that $|\lambda_k|$ are all convex. Set

$$\sigma(x) := A\,\mathrm{Diag}(\lambda_1(x), \ldots, \lambda_q(x))O,\ A \in \mathbb{M}_{d,q}(\mathbb{R}),\ O \in \mathcal{O}(q, \mathbb{R})$$

  then σ is $\preceq$-convex.

- When $q = d$, σ $\preceq$-convex is equivalent to

$$\sigma\sigma^*(\alpha x + (1-\alpha)y) \le \Big(\alpha\sqrt{\sigma\sigma^*}(x) + (1-\alpha)\sqrt{\sigma\sigma^*}(y)\Big)\Big(\alpha\sqrt{\sigma\sigma^*}(x) + (1-\alpha)\sqrt{\sigma\sigma^*}(y)\Big)^*.$$

Theorem (Strong martingale diffusion, P. 2016, Fadili-P. 2017, Jourdain-P. 2021)

Let $\sigma, \theta \in \mathrm{Lip}_x([0,T]\times\mathbb{R}^d, \mathbb{M}_{d,q})$, W q-S.B.M.. Let $X^{(\sigma)}$ and $X^{(\theta)}$ be the unique strong solutions to

$$\begin{aligned} dX_t^{(\sigma)} &= \sigma(t, X_t^{(\sigma)})dW_t, \; X_0^{(\sigma)} \in L^1(\mathbb{P}) \\ dX_t^{(\theta)} &= \theta(t, X_t^{(\theta)})dW_t, \; X_0^{(\theta)} \in L^1(\mathbb{P}), \quad (W_t)_{t\in[0,T]} \text{ standard B.M.} \end{aligned}$$

(a) *If $X_0^{(\sigma)} \preceq_{cvx} X_0^{(\theta)}$ and*

$$\begin{cases} (i)_\sigma & \sigma(t,\cdot) : \mathbb{R}^d \to \mathbb{M}_{d,q} \text{ is } \preceq\text{-convex for every } t \in [0,T], \; [|\sigma(t,\cdot)| \text{ convex}]_{d=q=1} \\ \text{or} & \\ (i)_\theta & \theta(t,\cdot) : \mathbb{R}^d \to \mathbb{M}_{d,q} \text{ is } \preceq\text{-convex for every } t \in [0,T], \; [|\theta(t,\cdot)| \text{ convex}]_{d=q=1} \\ \text{and} & \\ (ii) & \sigma(t,\cdot) \preceq \theta(t,\cdot) \text{ for every } t \in [0,T], \; [|\sigma(t,\cdot)| \le |\theta(t,\cdot)|]_{d=q=1} \end{cases}$$

then:

– for every l.s.c. convex $F : \mathcal{C}([0,T],\mathbb{R}^d) \to \mathbb{R}$, $\mathbb{E}\,F(X^{(\sigma)}) \le \mathbb{E}\,F(X^{(\theta)})$

– if $(i)_\sigma$ holds true, then one also have

$$x \mapsto \mathbb{E}\,F(X^{(\sigma),x}) \text{ is convex.}$$

- By a functional inf-convolution argument, it suffices to consider $\|\cdot\|_{\sup}$-Lipschitz functionals [Jourdain-P., Fin. & Stoch., '24].

Theorem (Weak Martingale diffusions, P. 2016, Fadili-P. 2017)

Let $\sigma, \theta \in \mathcal{C}_{lin_x,unif}\big([0,T]\times\mathbb{R}^d, \mathbb{M}_{d,q}\big)$, $W^{(\sigma)}$, $W^{(\theta)}$ *q-S.B.M.. Let* $X^{(\sigma)}$ *and* $X^{(\theta)}$ *be the unique weak solutions to*

$$\begin{aligned} dX_t^{(\sigma)} &= \sigma(t, X_t^{(\sigma)})dW_t^{(\sigma)},\ X_0^{(\sigma)} \in L^{1+\eta}(\mathbb{P}) \\ dX_t^{(\theta)} &= \theta(t, X_t^{(\theta)})dW_t^{(\theta)},\ X_0^{(\theta)} \in L^{1+\eta}(\mathbb{P}), \quad (W_t^{(\cdot)})_{t\in[0,T]} \text{ standard B.M.} \end{aligned}$$

(a) *If* $X_0^{(\sigma)} \preceq_{cvx} X_0^{(\theta)}$ *and*

$$\begin{cases} (i)_\sigma & \sigma(t,.) : \mathbb{R}^d \to \mathbb{M}_{d,q} \text{ is } \preceq\text{-convex for every } t \in [0,T],\ [|\sigma(t,\cdot)| \text{ convex}]_{d=q=1} \\ \text{or} & \\ (i)_\theta & \theta(t,.) : \mathbb{R}^d \to \mathbb{M}_{d,q} \text{ is } \preceq\text{-convex for every } t \in [0,T],\ [|\theta(t,\cdot)| \text{ convex}]_{d=q=1} \\ \text{and} & \\ (ii) & \sigma(t,\cdot) \preceq \theta(t,\cdot) \text{ for every } t \in [0,T],\ [|\sigma(t,\cdot)| \le |\theta(t,\cdot)|]_{d=q=1} \end{cases}$$

then:

– *for every l.s.c. convex* $F : \mathcal{C}([0,T],\mathbb{R}^d) \to \mathbb{R}$, $\mathbb{E}\,F(X^{(\sigma)}) \le \mathbb{E}\,F(X^{(\theta)})$

– *if* $(i)_\sigma$ *holds true and* F *has* $\|\,.\,\|_{\sup}$*-polynomial growth,*

$$x \mapsto \mathbb{E}\,F(X^{(\sigma),x}) \text{ is convex.}$$

The 1D case (martingale case)

Theorem (P. 2016)

Let $\sigma, \theta \in \mathcal{C}_{lin_x,unif}([0,T]\times\mathbb{R},\mathbb{R})$. *Let* $X^{(\sigma)}$ *and* $X^{(\theta)}$ *be unique weak solutions to*

$$dX_t^{(\sigma)} = \sigma(t, X_t^{(\sigma)})dW_t^{(\sigma)},\ X_0^{(\sigma)} \in L^1(\mathbb{P})$$
$$dX_t^{(\theta)} = \theta(t, X_t^{(\theta)})dW_t^{(\theta)},\ \ X_0^{(\theta)} \in L^1(\mathbb{P}),\ (W_t^{(\cdot)})_{t\in[0,T]} \text{ standard B.M.}$$

(a) *If* $X_0^{(\sigma)} \preceq_{cvx} X_0^{(\theta)}$ *and*

$$\begin{cases} (i)_\sigma & |\sigma(t,.)| : \mathbb{R} \to \mathbb{R}_+ \text{ is convex for every } t \in [0,T], \\ \text{or} & \\ (i)_\theta & |\theta(t,.)| : \mathbb{R} \to \mathbb{R}_+ \text{ is convex for every } t \in [0,T], \\ \text{and} & \\ (ii) & |\sigma(t,\cdot)| \le |\theta(t,\cdot)| \text{ for every } t \in [0,T] \end{cases}$$

then:

– *for every l.s.c. convex* $F : \mathcal{C}([0,T],\mathbb{R}^d) \to \mathbb{R}$, $\mathbb{E}\,F(X^{(\sigma)}) \le \mathbb{E}\,F(X^{(\theta)})$

– *if* $(i)_\sigma$ *holds true and* F *has* $\|\,.\,\|_{\sup}$*-polynomial growth*

$$x \mapsto \mathbb{E}\,F(X^{(\sigma),x}) \text{ is convex.}$$

Scaled/drifted martingale diffusions (extension to)

- The former theorems still hold true for

$$X_t^{(\sigma)} = X_0^{(\sigma)} + \int_0^t \alpha(t)\big(X_t^{(\sigma)} + \beta(t)\big)dt + \int_0^t \sigma(t, X_t^{(\sigma)})dW_t^{(\sigma)},$$
$$X_t^{(\theta)} = X_0^{(\theta)} + \int_0^t \alpha(t)\big(X_t^{(\theta)} + \beta(t)\big)dt + \int_0^t \theta(t, X_t^{(\theta)})dW_t^{(\theta)},$$

  where $\alpha(t) \in \mathbb{M}_{d,d}$ and $\beta(t) \in \mathbb{R}^d$ are Hölder continuous.

- Change of variable:

$$\widetilde{X}_t^{(\sigma)} = e^{-\int_0^t \alpha(s)ds}\big(X_t^{(\sigma)} + \beta(t)\big), \;\; \text{etc.}$$

- Finance: spot interest rate $\alpha(t) = r(t)\mathbf{1}$ and $\beta(t) = 0$ since typical (risk-neutral) dynamics of traded assets read

$$dS_t = r(t)S_t dt + S_t\sigma(S_t,)dW_t.$$

Functional Hajek's Theorem on Monotone convex ordering ($d = q = 1$)

Let

$$X_t^{(\sigma)} = X_0^{(\sigma)} + \int_0^t b_1(t, X_t^{(\sigma)})dt + \int_0^t \sigma(t, X_t^{(\sigma)})dW_t^{(\sigma)},$$

$$X_t^{(\theta)} = X_0^{(\theta)} + \int_0^t b_2(t, X_t^{(\theta)})dt + \int_0^t \theta(t, X_t^{(\theta)})dW_t^{(\theta)}.$$

where all coefficients $b_i(t,\cdot)$, $\sigma(t,\cdot)$, $\theta(t,\cdot)$ are Lipschitz, uniformly in $t \in [0,T]$.

Theorem (Strong solution version)

Assume furthermore

$$\begin{aligned} & (i)_\sigma \equiv b_1(t,\cdot) \text{ and } |\sigma(t,\cdot)| \text{ convex } \forall\, t \in [0,T]) \\ & \text{or} \\ & (i)_\theta \equiv b_2(t,\cdot) \text{ and } |\theta(t,\cdot)| \text{ convex } \forall\, t \in [0,T], \\ \text{and} \quad & (ii) \equiv \quad b_1(t,\cdot) \le b_2(t,\cdot)\ \&\ |\sigma(t,\cdot)| \le |\theta(t,\cdot)| \end{aligned}$$

and $X_0^{(\sigma)} \le_{icv} X_0^{(\theta)}$.

Theorem (continued)

Then:

– *for every l.s.c. convex, pointwise non-decreasing,* $F : \mathcal{C}([0,T],\mathbb{R}) \to \mathbb{R}$,

$$\mathbb{E}\, F(X^{(\sigma)}) \le \mathbb{E}\, F(X^{(\theta)}).$$

– *if* $(i)_\sigma$ *holds true*

$$x \mapsto \mathbb{E}\, F(X^{(\sigma),x}) \text{ is non-decreasing and convex.}$$

- Hajek's original theorem only dealt with marginal convex ordering.
- Assume $(*)_1$. One defines for f non-decreasing and convex and $0 < h < 1/[b_1]_{\rm Lip}$

$$Q_h f(x,u) = \mathbb{E}\, f\big(x + hb_1(x) + uZ\big)$$

which is convex and nondecreasing in both x and u.
- Mimic the (yet unknown!) proof of the previous theorem. □

Strategy (constructive)

- Time discretization (preferably) accessible to simulation: typically the Euler scheme.
- Propagate convexity (marginal or pathwise)
- Propagate comparison (marginal or pathwise)
- Transfer:

  – By $L^1(\mathbb{P})$ $\|\cdot\|_{\sup}$-convergence (hence $\|\cdot\|_{\sup}$ $\mathcal{W}_1$-convergence) of the Euler scheme (strong solutions setting)

  – by functional weak limit theorems "à la Jacod-Shiryaev" (weak solutions setting).

Step 1: discrete time ARCH models

- ARCH dynamics: Let $(Z_k)_{1\le k\le n}$ be a sequence of independent, radial r.v. on $(\Omega, \mathcal{A}, \mathbb{P})$. Two ARCH models: $X_0, Y_0 \in L^1(\mathbb{P})$,

$$\begin{aligned} X_{k+1} &= X_k + \sigma_k(X_k)\, Z_{k+1}, \\ Y_{k+1} &= Y_k + \theta_k(Y_k)\, Z_{k+1}, \quad k = 0 : n-1, \end{aligned}$$

where σ_k, $\theta_k : \mathbb{R} \to \mathbb{R}$, $k = 0 : n-1$ have linear growth.

Proposition (Propagation result)

If σ_k, $k = 0 = n-1$ are $\preceq$-convex with linear growth,

$$X_0 = x \quad \text{and} \quad \forall\, k \in \{0, \ldots, n-1\}, \quad \sigma_k \preceq \theta_k,$$

then, for every convex function $F : (\mathbb{R}^d)^{n+1} \to \mathbb{R}$ convex with linear growth

$$x \longmapsto \mathbb{E}\, F(x, X_1^x \ldots, X_n^x) \quad \text{is convex.}$$

Partial proof (marginal) with radial white noise

- $Z_k \sim \mathcal{N}(0, I_q)$, $1 \leq k \leq n$ or, more generally, $Z_k \sim OZ_k$, $\forall\, O \in \mathcal{O}(q, \mathbb{R})$.
- Let $f : \mathbb{R}^d \to \mathbb{R}$ be a convex function (with linear growth). Let

$$P_k^\sigma f(x) := \mathbb{E}\, f(x + \sigma_{k-1}(x) Z_k) = \big[\mathbb{E}\, f(x + AZ_k)\big]_{|A=\sigma_{k-1}(x)}.$$

- Let $\mathbb{R}^d \times \mathbb{M}_{d,q} \ni (x, A) \mapsto Q_k f(x, A) := \mathbb{E}\, f(x + AZ_k)$, $k = 1 : n$. It is convex, right $O(q, \mathbb{R})$-invariant and $\preceq$-non-decreasing in A by the starting example.
  - $Q_k f(x, AO) = \mathbb{E}\, f(x + AOZ_k) = \mathbb{E}\, f(x + AZ_k)$,
  - $Q_k f(\lambda x + (1-\lambda) y, \lambda A + (1-\lambda) B)) = \mathbb{E}\, f\big(\lambda(x + AZ_k) + (1-\lambda)(y + BZ_k)\big)$
    $\leq \lambda Q_k f(x, A) + (1-\lambda) Q_k f(y, B)$ by convexity of f.
  - If $A \preceq B$, then $AZ_k \preceq_{cvx} BZ_k$ and $f(x + \cdot)$ is convex.
- Hence if x, $y \in \mathbb{R}^d$ and $\lambda \in [0, 1]$

$$\begin{aligned} P_k^\sigma f\big(\lambda x + (1-\lambda) y\big) &= Q_k f\big(\lambda x + (1-\lambda) y, \sigma_{k-1}(\lambda x + (1-\lambda) y)\big) \\ &\leq Q_k f\big(\lambda x + (1-\lambda) y, \lambda \sigma_{k-1}(x) + (1-\lambda) \sigma_{k-1}(y)\big) \\ Q_k \text{ convex and } \geq 0 \text{ operator} \longrightarrow\quad &\leq \lambda Q_k f\big(x, \sigma_{k-1}(x)\big) + (1-\lambda) Q_k f\big(y, \sigma_{k-1}(y)\big) \\ &= \lambda P_k^\sigma f(x) + (1-\lambda) P_k^\sigma f(y). \end{aligned}$$

- Hence the transition kernels P_k^{σ} propagate convexity:

$$f \;\text{ convex} \Longrightarrow P_k^{\sigma}(f) \;\text{ convex.}$$

- by an either forward or backward induction on k, one finally gets.

$$x \longmapsto \mathbb{E}\, f(X_n^x) = P_{1:n}^{\sigma} f(x) := P_1^{\sigma} \circ \cdots \circ P_n^{\sigma} f(x) \qquad \text{is convex.}$$

Proposition (Discrete time convex ordering result)

If all σ_k, $k = 0 : n-1$ or all θ_k, $k = 0 : n-1$ are $\preceq$-convex with (at most) linear growth,

$$X_0 \preceq_{cvx} Y_0 \quad \text{and} \quad \forall\, k \in \{0, \ldots, n-1\}, \quad \sigma_k \preceq \theta_k,$$

then

$$(X_0, \ldots, X_n) \preceq_{cvx} (Y_0, \ldots, Y_n).$$

Partial proof (marginal) with radial white noise

- Assume e.g. that all σ_k are convex.
- Backward induction on k.
- For $k = n$. Let $f : \mathbb{R}^d \to \mathbb{R}$ be a convex function with linear growth.

$$P_n^{\sigma} f(x) = Q_n f\big(x, \sigma_{n-1}(x)\big) \le Q_n f\big(x, \theta_{n-1}(x)\big) = P_n^{\theta} f(x)$$

  by non-decreasing $\preceq$-monotony of Q_n.

- Assume $\underbrace{P^{\sigma}_{k+1:n} f}_{\text{convex}} \le P^{\theta}_{k+1:n} f$. Then

$$\forall\, x \in \mathbb{R}^d, \quad \mathbb{M}_{d,q} \ni A \longmapsto Q_k\big(P^{\sigma}_{k+1:n} f\big)(x, A) \quad \text{is } \preceq\text{-non-decreasing}$$

so that

$$\begin{aligned} P^{\sigma}_{k:n} f(x) &= Q_k\big(P^{\sigma}_{k+1:n} f\big)\big(x, \sigma_{k-1}(x)\big) \overset{\downarrow}{\le} Q_k\big(P^{\sigma}_{k+1:n} f\big)\big(x, \theta_{k-1}(x)\big) \\ Q_k \text{ positive operator } \longrightarrow &\le Q_k\big(P^{\theta}_{k+1:n} f\big)\big(x, \theta_{k-1}(x)\big) \\ &= P^{x,\theta}_{k:n} f(x). \end{aligned}$$

- Hence, in particular for $f : \mathbb{R}^d \to \mathbb{R}$ Lipschitz and convex

$$\mathbb{E}\, f(X_n^{\sigma}) = \mathbb{E}\, P^{\sigma}_{1:n} f(X_0) \le \mathbb{E}\, P^{\sigma}_{1:n} f(Y_0) \le \mathbb{E}\, P^{\theta}_{1:n} f(Y_0) = \mathbb{E}\, f(X_n^{\theta}). \qquad \square$$

Global convex ordering

- Same strategy
- But entirely backward.
- $q = d = 1$ for simplicity.

▷ Dynamic programming: We introduce two martingales

$$M_k = \mathbb{E}\big(F(X_{0:n})\,|\,\mathcal{F}_k^Z\big) \quad\text{and}\quad N_k = \mathbb{E}\big(F(Y_{0:n})\,|\,\mathcal{F}_k^Z\big),\; k = 0:n$$

and again the sequence of operators

$$Q_k(f)(x,u) = \mathbb{E}\,f(x+uZ_k),\; u\in\mathbb{R},\; k = 1:n.$$

Warning (for the mini-course)

- For convenience we will make the proof in a one-dimensional setting.
- Then a slightly revisited version of Jensen's inequality simplifies the communication.
- It follows ([4])

[4] G. Pagès, Convex order for path-dependent derivatives: a dynamic programing approach, Séminaire de Probabilités, XLVIII, LNM 2168, Springer, Berlin, 33-96, 2016.

Jensen's Inequality (a bit) revisited = Key Lemma

Lemma (Jensen's Inequality revisited)

Let $Z : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{R}$ *be an centered integrable r.v.:* $Z \in L^1(\mathbb{P})$, $\mathbb{E}\, Z = 0$.

▷ *Let* $f : \mathbb{R} \to \mathbb{R}$, *convex, such that*

$$\forall\, x, u \in \mathbb{R},\;\; Qf(x,u) := \mathbb{E}\, f\big(x + u\, Z\big) \text{ is well-defined in } \mathbb{R}.$$

Then $Qf(x,\cdot)$ *is convex, attains its minimum at* 0 *so that*

$$Qf(x,.) \text{ is non-decreasing on } \mathbb{R}_+, \text{ non-increasing on } \mathbb{R}_-.$$

▷ *If* $Z \sim -Z$ *(symmetric distribution), then* $Qf(x+\cdot)$ *is an even function and*

$$\forall\, x \in \mathbb{R},\; \forall\, a \in \mathbb{R}_+, \quad \sup_{|u|\le a} Qf(x,u) = Qf(x,a).$$

Proof. The function Qf is clearly convex and by Jensen's Inequality

$$Qf(x,u) \ge f\big(\mathbb{E}(x + u\, Z)\big) = f\big(x + u\, \mathbb{E}\, Z\big) = f(x) = Qf(x,0).$$

Hence $Q\,f$ is convex, $Qf(x+\cdot)$ attains its minimum at $u = 0$ hence is non-increasing on $\mathbb{R}_-$ and non-decreasing on $\mathbb{R}_+$. □

- A (first) backward induction and the definition of the kernels Q_k imply

$$M_k = \Phi_k(X_{0:k}) \quad \text{and} \quad N_k = \Psi_k(Y_{0:k}), \quad k = 0, \ldots, n.$$

where $\Phi_k, \Psi_k : \mathbb{R}^{k+1} \to \mathbb{R}$, $k = 0, \ldots, n$ are recursively defined by

$$\begin{aligned}\Phi_n &:= F,\\ \Phi_k(x_{0:k}) &= \left[\mathbb{E}\, \Phi_{k+1}(x_{0:k}, x_k + uZ_{k+1})\right]_{|u=\sigma_k(x_k)}\\ &:= \big(Q_{k+1}\Phi_{k+1}(x_{0:k}, \cdot)\big)(x_k, \sigma_k(x_k)), \quad k = 0 : n-1.\end{aligned}$$

Likewise

$$\Psi_n := F, \;\; \Psi_k(y_{0:k}) := \big(Q_{k+1}\Psi_{k+1}(y_{0:k}, \cdot)\big)(y_k, \theta_k(y_k)), \; k = 0 : n-1.$$

▷ Assume now that all functions σ_k are ≥ 0 and convex:

Lemma

$$\Big(G : \mathbb{R}^{k+2} \to \mathbb{R} \text{ convex}\Big)$$
$$\Downarrow$$
$$\Big((x_{0:k}, u) \mapsto \mathbb{E}\,G(x_{0:k}, x_k + uZ_{k+1}) = Q_{k+1}G(x_{0:k}, \cdot)(x_k, u) \text{ is convex}\ldots\Big)$$

so that, by the revisited Jensen's Lemma, one has

(*i*) $u \mapsto Q_{k+1}G(x_{0:k}, \cdot)(x_k, u)$ *is* $\downarrow$ *on* $(-\infty, 0)$ *and* $\uparrow$ *on* $(0, +\infty)$.

&

(*ii*) *Propagation of the convexity in* $x_{0:k}$.

▷ Assume now that all functions σ_k are ≥ 0 and convex:

Lemma

$$\Big(G:\mathbb{R}^{k+2}\to\mathbb{R} \textit{ convex}\Big)$$
$$\Downarrow$$
$$\Big((x_{0:k},u)\mapsto \mathbb{E}\,G(x_{0:k},x_k+uZ_{k+1}) = Q_{k+1}G(x_{0:k},\cdot)(x_k,u) \textit{ is convex}\ldots\Big)$$

so that, by the revisited Jensen's Lemma, one has

(*i*) $u\mapsto Q_{k+1}G(x_{0:k},\cdot)(x_k,u)$ *is* $\downarrow$ *on* $(-\infty,0)$ *and* $\uparrow$ *on* $(0,+\infty)$.

&

(*ii*) *Propagation of the convexity in* $x_{0:k}$.

- (Second) backward induction $\Longrightarrow$ all functions Φ_k are convex.

- (Third) backward induction $\Longrightarrow \Phi_k \leq \Psi_k,\ k = 0 : n-1$.

  First note that $\Phi_n = \Psi_n = F$. If $\Phi_{k+1} \leq \Psi_{k+1}$, then

$$
\begin{aligned}
\Phi_k(x_{0:k}) &= \big(Q_{k+1}\Phi_{k+1}(x_{0,k},\cdot)\big)(x_k, \sigma_k(x_k))\\
&\leq \big(Q_{k+1}\Phi_{k+1}(x_{0:k},\cdot)\big)(x_k, \theta_k(x_k))\\
&\leq \big(Q_{k+1}\Psi_{k+1}(x_{0:k},\cdot)\big)(x_k, \theta_k(x_k)) = \Psi_k(x_{0:k}).
\end{aligned}
$$

- When $k = 0$

$$\Phi_0 \text{ convex and } \Phi_0(x) \leq \Psi_0(x) \quad \Longleftrightarrow \quad \mathbb{E}\, F(X_{0:n}) \leq \mathbb{E}\, F(Y_{0:n}).$$

  so that

$$\mathbb{E}\, F(X_{0:n}) = \mathbb{E}\, \Phi_0(X_0) \leq \mathbb{E}\, \Phi_0(Y_0) \leq \mathbb{E}\, \Psi_0(Y_0) = \mathbb{E}\, F(Y_{0:n}).$$

End of discrete time setting

▷ If all $\theta_k \geq 0$ and convex:

This time, one shows that:

- the functions Ψ_k are convex, $k = 0, \ldots, n$
- $\Phi_n \leq \Psi_n \Longrightarrow \Phi_k \leq \Psi_k$, $k = 0, \ldots, n-1$.

Remark. The discrete time setting has its own interest.

Step 2 of the proof: Back to continuous time

▷ Euler scheme(s): Discrete time Euler scheme with step $\frac{T}{n}$, starting at x is an ARCH model. For $X^{(\sigma)}$: for $k = 0, \ldots, n-1$,

$$\bar{X}^{(\sigma),n}_{t^n_{k+1}} = \bar{X}^{(\sigma),n}_{t^n_k} + \sigma\big(t^n_k, \bar{X}^{(\sigma),n}_{t^n_k}\big)\big(W_{t^n_{k+1}} - W_{t^n_k}\big), \; \bar{X}^{(\sigma),n}_0 = x$$

Set

$$Z_k = W_{t^n_k} - W_{t^n_{k-1}}, \; k = 1, \ldots, n, \; i.i.d.$$

$$\Downarrow$$

discrete time setting applies

Remark. Linear growth of σ and θ, implies if $X^{(\sigma)}_0$ and $X^{(\theta)}_0 \in L^p(\mathbb{P})$ for $p = 1 + \eta > 1$, then

$$\sup_{n\geq 1}\Big\|\sup_{t\in[0,T]}|\bar{X}^{(\sigma),n}_t|\Big\|_p + \sup_{n\geq 1}\Big\|\sup_{t\in[0,T]}|\bar{X}^{(\theta),n}_t|\Big\|_p \leq C(1 + \|X_0\|_p).$$

From discrete to continuous time

▷ Interpolation $(n \geq 1)$

- *Piecewise affine interpolator* defined by

$$\forall\, x_{0:n} \in \mathbb{R}^{n+1},\ \forall\, k = 0, \ldots, n-1,\ \forall\, t \in [t_k^n, t_{k+1}^n], \quad .$$

$$i_n(x_{0:n})(t) = \frac{n}{T}\Big((t_{k+1}^n - t)x_k + (t - t_k^n)x_{k+1}\Big)$$

- $\widetilde{X}^{(\sigma),n} := i_n\Big((\bar{X}^{(\sigma),n}_{t_k^n})_{k=0:n}\Big) =$ piecewise affine Euler scheme.

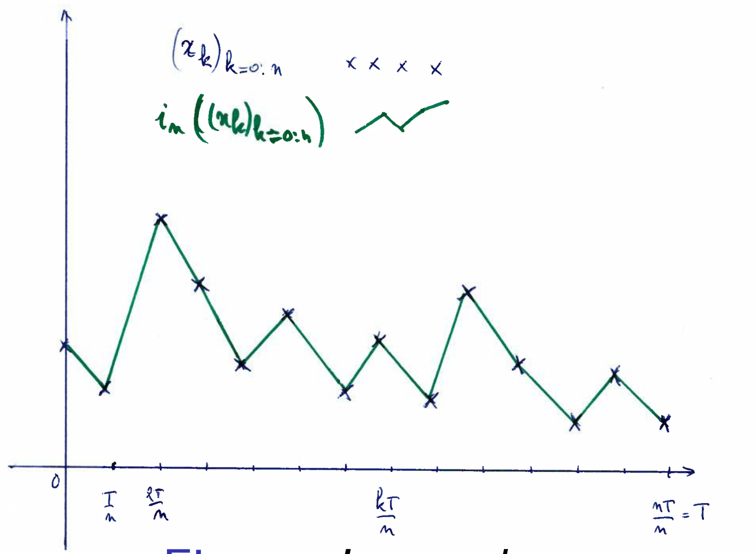

Figure: *Interpolator*

"Strong" solution setting

▷ Let $F : \mathcal{C}([0,T],\mathbb{R}) \to \mathbb{R}$ be a Lipschitz convex functional.

$$\forall\, n \geq 1, \qquad F_n : \mathbb{R}^{n+1} \ni x_{0:n} \longmapsto F_n(x_{0:n}) := F\big(i_n(x_{0:n})\big).$$

- Step 1 (Discrete time): $F\big(\widetilde{X}^{(\sigma),n}\big) = F_n\big((\bar{X}^{(\sigma),n}_{t_k^n})_{k=0:n}\big)$ and

$$F \text{ convex} \Longrightarrow F_n \text{ convex}, \;\; n \geq 1.$$

Discrete time result implies, since $\sigma(t_k^n,.) \preceq \theta(t_k^n,.)$,

$$\mathbb{E}\, F\big(\widetilde{X}^{(\sigma),n}\big) = \mathbb{E}\, F_n\big((\bar{X}^{(\sigma),n}_{t_k^n})_{k=0:n}\big) \leq \mathbb{E}\, F_n\big((\bar{X}^{(\theta),n}_{t_k^n})_{k=0:n}\big) = \mathbb{E}\, F\big(\widetilde{X}^{(\theta),n}\big).$$

- Step 2 (Transfer in the "strong" Lipschitz setting): We know that

$$\mathcal{W}_1\big(\widetilde{X}^{(\sigma),n}, X^{(\sigma)}\big) \leq \Big\| \big\|\widetilde{X}^{(\sigma),n} - X^{(\sigma)}\big\|_{\sup} \Big\|_1 \to 0 \quad \text{as } n \to +\infty$$

Hence if $F : \mathcal{C}([0,T],\mathbb{R}^d) \to \mathbb{R}$ is $\|\cdot\|_{\sup}$-Lipschitz

$$\Big|\mathbb{E}\, F\big(\widetilde{X}^{(\sigma),n}\big) - \mathbb{E}\, F\big(X^{(\sigma)}\big)\Big| \leq [F]_{\rm Lip}\mathcal{W}_1\big(\widetilde{X}^{(\sigma),n}, X^{(\sigma)}\big) \to 0 \quad \text{as } n \to +\infty$$

Idem for the $X^{(\theta)}$-diffusion, so that

$$\mathbb{E}\, F(X^{(\sigma)}) \leq \mathbb{E}\, F(X^{(\theta)}). \qquad \square$$

"Weak" diffusion setting

- Step 2bis (Transfer in the "weak" linear growth continuous setting): See e.g. [Jacod-Shiryaev's book 2^{nd} edition, Theorem 3.39, p.551] ([5]).

$$\widetilde{X}^{(\sigma),n} \stackrel{\mathcal{L}(\|.\|_{\sup})}{\longrightarrow} X^{(\sigma)} \quad \text{and} \quad \widetilde{X}^{(\sigma),n} \stackrel{\mathcal{L}(\|.\|_{\sup})}{\longrightarrow} X^{(\theta)} \quad \text{as } n \to +\infty.$$

- We know that, as $\sigma(t,\cdot)$ and $\theta(t,\cdot)$ have linear growth

$$\Big\| \sup_{t\in[0,T]} |\widetilde{X}^{(\sigma),n}| \Big\|_{1+\eta} + \Big\| \sup_{t\in[0,T]} |\widetilde{X}^{(\theta),n}| \Big\|_{1+\eta} \le C_{\eta,T}(1 + \|X_0\|_{1+\eta})$$

Hence, if F is $\|\cdot\|_{\sup}$-Lipschitz, then $F\big(\widetilde{X}^{(\sigma),n}\big)$, $n \ge 1$, is uniformly integrable so that

$$\mathbb{E}\, F(X^{(\sigma)}) = \lim_n \mathbb{E}\, F\big(\widetilde{X}^{(\sigma),n}\big) \qquad (\text{idem for } X^{(\theta)}).$$

- Hence $\quad \mathbb{E}\, F(X^{(\sigma)}) \le \mathbb{E}\, F(X^{(\theta)}).$ □

[5] *Limit theorems for stochastic processes*, Springer, 2010.