# Newton method for stochastic control problems

Emmanuel GOBET -Maxime GRANGEREAU

CMAP, Ecole polytechnique and Osiris, EDF R&D

To appear in *SIAM Control and Optimization*

BSDE 2022 at Annecy
June 27-July 1 2022

# Introduction: goals

- Newton method:
  - Iterative method.
  - Linearized critical point equation yields a new search direction (Newton step)
  - Globalization: choose appropriate step length (line-search).
  - Expected to converge faster than naive iterative method.
- Our purpose:
  - design this principle for solving stochastic control.
  - Establish global convergence and quick local convergence.
  - Practical implementation.

1. Recall on Newton method for optimization in $\mathbb{R}^d$
2. Stochastic control: framework and Newton step computation
3. Theoretical convergence properties
4. Numerical implementation and results

1. Recall on Newton method for optimization in $\mathbb{R}^d$
2. Stochastic control: framework and Newton step computation
3. Theoretical convergence properties
4. Numerical implementation and results

# Newton method in $\mathbb{R}^d$: principle for $f \in C^2(\mathbb{R}^d, \mathbb{R})$

- Second-order iterative method for optimization problems

$$\min_{x \in \mathbb{R}^d} f(x).$$

- First order optimality condition (sufficient if $f$ is convex)

$$\nabla f(x^*) = 0.$$

# Newton method in $\mathbb{R}^d$: principle for $f \in C^2(\mathbb{R}^d, \mathbb{R})$

- Second-order iterative method for optimization problems

$$\min_{x \in \mathbb{R}^d} f(x).$$

- First order optimality condition (sufficient if $f$ is convex)

$$\nabla f(x^*) = 0.$$

- Successive linear approximations of the critical point equation

$$\nabla f(x^{(k)}) + \nabla^2 f(x^{(k)})\ \Delta_x = 0.$$

- New iterate

$$x^{(k+1)} = x^{(k)} + \Delta_x.$$

# Newton method in $\mathbb{R}^d$: convergence properties

If $f$ has a unique minimizer $x^*$ and is:

- twice differentiable with Lipschitz second-order derivatives,
- strongly convex

$$\exists \alpha > 0, \forall (x, y), \quad f(y) \geq f(x) + \nabla f(x) \cdot (y - x) + \frac{\alpha}{2} \|y - x\|^2.$$

Then $(x^{(k)})_{k \in \mathbb{N}}$ converges quadratically locally to $x^*$ [3]

$$x^{(0)} \in V \Rightarrow \forall k \in \mathbb{N}, C\|x^{(k+1)} - x^*\| \leq \left(C\|x^{(k)} - x^*\|\right)^2.$$

Newton method converges in 1 iteration if $f$ is quadratic

# Illustration of local convergence of Newton method

Strongly convex function $f$ with Lipschitz second-order derivative:

$$f : x \mapsto |x|^3 + x^2$$

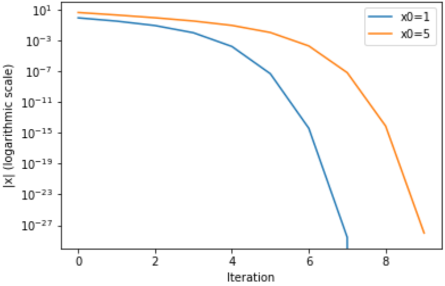


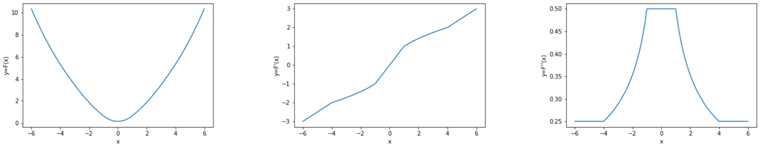

Figure: Distance of current iterate $x^k$ to minimizer $x^* = 0$

Observe 2 convergence regimes.

# Newton method may not converge globally (in $\mathbb{R}^d$)

Strongly convex function $f$ with Lipschitz second-order derivative:

$$f(x) := \begin{cases} \frac{x^2}{4} + \frac{4}{3}, & \text{if } |x| > 4, \\ \frac{2|x|^{\frac{3}{2}}}{3}, & \text{if } 1 \leq |x| \leq 4, \\ \frac{x^2}{2} + \frac{1}{6}, & \text{if } |x| < 1. \end{cases}$$




(a) Graph of $f$ (b) Graph of $f'$ (c) Graph of $f''$

If $|x^{(0)}| \in (1,4)$, then $x^{(k+1)} = x^{(k)} - \frac{f'(x^{(k)})}{f''(x^{(k)})} = -x^{(k)}$.
No convergence, need to choose a stepsize.

# Newton method with backtracking line-search (in $\mathbb{R}^d$)

- Take step $\sigma\Delta_x$ instead of full Newton step $\Delta_x$...
- ... with $\sigma$ largest value in $\{1, \beta, \beta^2, \beta^3...\}$ such that

$$f(x^{(k)} + \sigma\Delta_x) \leq f(x^{(k)}) + \gamma\sigma\nabla f(x^{(k)})\ \Delta_x.$$

# Newton method with backtracking line-search (in $\mathbb{R}^d$)

- Take step $\sigma\Delta_x$ instead of full Newton step $\Delta_x$...
- ... with $\sigma$ largest value in $\{1, \beta, \beta^2, \beta^3 ...\}$ such that

$$f(x^{(k)} + \sigma\Delta_x) \leq f(x^{(k)}) + \gamma\sigma\nabla f(x^{(k)})\ \Delta_x.$$

**Theorem**

*There is a $k^*$ s.t.*

- *Until $k^*$: arithmetic convergence*

$$\exists \eta > 0, \forall k \leq k^*, \quad \|x^{(k+1)} - x^*\| \leq \|x^{(k)} - x^*\| - \eta.$$

- *After $k^*$: quadratic convergence*

$$\exists C > 0, \forall k > k^*, \quad C\|x^{(k+1)} - x^*\| \leq \left(C\|x^{(k)} - x^*\|\right)^2.$$

1 Recall on Newton method for optimization in $\mathbb{R}^d$

2 Stochastic control: framework and Newton step computation

3 Theoretical convergence properties

4 Numerical implementation and results

# Mathematical formulation

- $u$: control variable (prog. mes. square integrable process)
- $X^u$: state of system (ODE controlled by $u$)
- Unconstrained convex problem: **random** linear dynamic (parameter $\alpha$), random convex running cost function $l$ and terminal cost function $\Psi$.

$$\min_u \mathcal{J}(u) := \mathbb{E}\left[\int_0^T l(t, u_t, X_t^u)\,\mathrm{d}t + \Psi(X_T^u)\right]$$

$$s.t. \quad X_t^u := x_0 + \int_0^t \alpha_s u_s \mathrm{d}s.$$

- We focus on one-dimensional control/state processes.
- Usual regularity and growth conditions on $l$ and $\Psi$.

# Stochastic Pontryagine principle

## Theorem

*Under regularity assumptions, for any* $u \in \mathbb{H}^2$, *define* $Y^u \in \mathbb{H}^{\infty,2}$ *by:*

$$Y_t^u = \mathbb{E}_t\left[\Psi'_x(X_T^u) + \int_t^T l'_x(s, u_s, X_s^u)\mathrm{d}s\right].$$

*Besides,* $\mathcal{J}$ *is Fréchet-differentiable with gradient at u denoted* $\nabla\mathcal{J}(u) \in \mathbb{H}^2$ *given by:*

$$(\nabla\mathcal{J}(u))_t = l'_u(t, u_t, X_t^u) + \alpha_t Y_t^u,$$
$$\|\nabla\mathcal{J}(u) - \mathcal{J}(v)\|_{\mathbb{H}^2} \leq C\|u - v\|_{\mathbb{H}^2}.$$

# Stochastic Pontryagine principle

## Theorem

*Under regularity assumptions, for any $u \in \mathbb{H}^2$, define $Y^u \in \mathbb{H}^{\infty,2}$ by:*

$$Y_t^u = \mathbb{E}_t\left[\Psi'_x(X_T^u) + \int_t^T l'_x(s, u_s, X_s^u)\mathrm{d}s\right].$$

*Besides, $\mathcal{J}$ is Fréchet-differentiable with gradient at $u$ denoted $\nabla\mathcal{J}(u) \in \mathbb{H}^2$ given by:*

$$(\nabla\mathcal{J}(u))_t = l'_u(t, u_t, X_t^u) + \alpha_t Y_t^u,$$
$$\|\nabla\mathcal{J}(u) - \mathcal{J}(v)\|_{\mathbb{H}^2} \leq C\|u - v\|_{\mathbb{H}^2}.$$

*Under additional convexity assumptions, $\nabla\mathcal{J}(u) = 0$ characterizes the (unique) optimal control.*

# Characterization of Newton step

## Theorem

▷ *The mapping* $\Phi_Y : u \in \mathbb{H}^2 \mapsto Y^u \in \mathbb{H}^{\infty,2}$ *is Gateaux-differentiable, and its Gateaux derivative at* $u$ *in direction* $v$ *is given by* $D\Phi_Y(u)(v) = \dot{Y}^{u,v}$*, solution of the (linear) BSDE:*

$$\dot{Y}_t^{u,v} = \mathbb{E}_t\left[\Psi''_{xx}(X_T^u)\dot{X}_T^v + \int_t^T \left(l''_{xu}(s, u_s, X_s^u)v_s + l''_{xx}(s, u_s, X_s^u)\dot{X}_s^v\right)\mathrm{d}s\right].$$

*Besides, we have* $\|\dot{Y}^{u,v}\|_{\mathbb{H}^{\infty,2}} \leq C\|v\|_{\mathbb{H}^2}$.

# Characterization of Newton step

## Theorem

▷ *The mapping $\Phi_Y : u \in \mathbb{H}^2 \mapsto Y^u \in \mathbb{H}^{\infty,2}$ is Gateaux-differentiable, and its Gateaux derivative at u in direction v is given by $D\Phi_Y(u)(v) = \dot{Y}^{u,v}$, solution of the (linear) BSDE:*

$$\dot{Y}_t^{u,v} = \mathbb{E}_t\left[\Psi''_{xx}(X_T^u)\dot{X}_T^v + \int_t^T \left(l''_{xu}(s, u_s, X_s^u)v_s + l''_{xx}(s, u_s, X_s^u)\dot{X}_s^v\right) \mathrm{d}s\right].$$

*Besides, we have $\|\dot{Y}^{u,v}\|_{\mathbb{H}^{\infty,2}} \leq C\|v\|_{\mathbb{H}^2}$.*

▷ *$\mathcal{J}$ is twice Gateaux-differentiable and $\nabla^2\mathcal{J} : \mathbb{H}^2 \mapsto \mathcal{L}(\mathbb{H}^2)$ is:*

$$\left(\nabla^2\mathcal{J}(u)(v)\right)_t = l''_{uu}(t, u_t, X_t^u)v_t + l''_{ux}(t, u_t, X_t^u)\dot{X}_t^v + \alpha_t\dot{Y}_t^{u,v}.$$

*Besides, $\nabla^2\mathcal{J}(u) : \mathbb{H}^2 \mapsto \mathbb{H}^2$ is a continuous endomorphism ($\|\nabla^2\mathcal{J}(u)(v)\|_{\mathbb{H}^2} \leq C\|v\|_{\mathbb{H}^2}$)*

Newton step. We aim at computing $\Delta u$

$$\nabla^2 \mathcal{J}(u)(\Delta u) = -\nabla \mathcal{J}(u).$$

Crucial interpretation in terms of LQ stochastic control problem (explicitly tractable using Riccati equations and LBSDE [14, 2]).

## Theorem

*Let* $(u, w) \in \mathbb{H}^2 \times \mathbb{H}^2$. *Consider* $\min_{v \in \mathbb{H}^2} \tilde{\mathcal{J}}^{LQ,u,w}(v)$ *s.t.* $\tilde{X}_t = \int_0^t \alpha_s v_s \mathrm{d}s$, *where* $\tilde{\mathcal{J}}^{LQ,u,w}(v)$ *is defined by:*

$$\mathbb{E}\left[\int_0^T \left\{\frac{1}{2} l''_{uu}(t, u_t, X_t^u)\, v_t^2 + \frac{1}{2} l''_{xx}(t, u_t, X_t^u)\, \tilde{X}_t^2 + l''_{ux}(t, u_t, X_t^u)\, \tilde{X}_t v_t - w_t v_t\right\} \mathrm{d}t + \frac{1}{2} \Psi''_{xx}(X_T^u) \tilde{X}_T^2\right].$$

Newton step. We aim at computing $\Delta u$

$$\nabla^2 \mathcal{J}(u)(\Delta u) = -\nabla \mathcal{J}(u).$$

Crucial interpretation in terms of LQ stochastic control problem (explicitly tractable using Riccati equations and LBSDE [14, 2]).

### Theorem

*Let* $(u, w) \in \mathbb{H}^2 \times \mathbb{H}^2$*. Consider* $\min_{v \in \mathbb{H}^2} \tilde{\mathcal{J}}^{LQ,u,w}(v)$ *s.t.* $\tilde{X}_t = \int_0^t \alpha_s v_s \mathrm{d}s$*, where* $\tilde{\mathcal{J}}^{LQ,u,w}(v)$ *is defined by:*

$$\mathbb{E}\left[\int_0^T \left\{\frac{1}{2} l''_{uu}(t, u_t, X_t^u)\, v_t^2 + \frac{1}{2} l''_{xx}(t, u_t, X_t^u)\, \tilde{X}_t^2 + l''_{ux}(t, u_t, X_t^u)\, \tilde{X}_t v_t - w_t v_t\right\} \mathrm{d}t + \frac{1}{2}\Psi''_{xx}(X_T^u)\tilde{X}_T^2\right].$$

*Then* $\tilde{\mathcal{J}}^{LQ,u,w}$ *has a unique minimizer* $\tilde{u}^{u,w} \in \mathbb{H}^2$ *characterized by:*

$$\begin{cases} \tilde{X}_t^{u,w} = \int_0^t \alpha_s \tilde{u}_s^{u,w} \mathrm{d}s, \\ \tilde{Y}_t^{u,w} = \mathbb{E}_t\left[\Psi''_{xx}(X_T^u)\tilde{X}_T^{u,w} + \int_t^T \left(l''_{xu}(s, u_s, X_s^u)\tilde{u}_s^{u,w} + l''_{xx}(s, u_s, X_s^u)\tilde{X}_s^{u,w}\right) \mathrm{d}s\right], \\ l''_{uu}(t, u_t, X_t^u)\tilde{u}_t^{u,w} + l''_{ux}(t, u_t, X_t^u)\tilde{X}_t^{u,w} + \alpha_t \tilde{Y}_t^{u,w} = w_t. \end{cases}$$

Newton step. We aim at computing $\Delta u$

$$\nabla^2 \mathcal{J}(u)(\Delta u) = -\nabla \mathcal{J}(u).$$

Crucial interpretation in terms of LQ stochastic control problem (explicitly tractable using Riccati equations and LBSDE [14, 2]).

**Theorem**

*Let $(u, w) \in \mathbb{H}^2 \times \mathbb{H}^2$. Consider* $\min_{v \in \mathbb{H}^2} \tilde{\mathcal{J}}^{LQ,u,w}(v)$ *s.t.* $\tilde{X}_t = \int_0^t \alpha_s v_s \mathrm{d}s$, *where* $\tilde{\mathcal{J}}^{LQ,u,w}(v)$ *is defined by:*

$$\mathbb{E}\left[\int_0^T \left\{\frac{1}{2} l''_{uu}(t, u_t, X_t^u)\, v_t^2 + \frac{1}{2} l''_{xx}(t, u_t, X_t^u)\, \tilde{X}_t^2 + l''_{ux}(t, u_t, X_t^u)\, \tilde{X}_t v_t - w_t v_t\right\} \mathrm{d}t + \frac{1}{2}\Psi''_{xx}(X_T^u)\tilde{X}_T^2\right].$$

*Then* $\tilde{\mathcal{J}}^{LQ,u,w}$ *has a unique minimizer* $\tilde{u}^{u,w} \in \mathbb{H}^2$ *characterized by:*

$$\begin{cases} \tilde{X}_t^{u,w} = \int_0^t \alpha_s \tilde{u}_s^{u,w} \mathrm{d}s, \\ \tilde{Y}_t^{u,w} = \mathbb{E}_t\left[\Psi''_{xx}(X_T^u)\tilde{X}_T^{u,w} + \int_t^T \left(l''_{xu}(s, u_s, X_s^u)\tilde{u}_s^{u,w} + l''_{xx}(s, u_s, X_s^u)\tilde{X}_s^{u,w}\right)\mathrm{d}s\right], \\ l''_{uu}(t, u_t, X_t^u)\tilde{u}_t^{u,w} + l''_{ux}(t, u_t, X_t^u)\tilde{X}_t^{u,w} + \alpha_t \tilde{Y}_t^{u,w} = w_t. \end{cases}$$

*Besides, for any* $u \in \mathbb{H}^2$, $\nabla^2 \mathcal{J}(u) \in \mathcal{L}(\mathbb{H}^2)$ *is invertible and* $\left(\nabla^2 \mathcal{J}(u)\right)^{-1}(w) = \tilde{u}^{u,w}$.

1. Recall on Newton method for optimization in $\mathbb{R}^d$
2. Stochastic control: framework and Newton step computation
3. Theoretical convergence properties
4. Numerical implementation and results

# Insufficient regularity properties in $\mathbb{H}^2$

Problem:

$$\min_{u \in \mathbb{H}^2} \mathcal{J}(u)$$

$\nabla^2 \mathcal{J}$: Lipschitz-continuous $\Rightarrow$ local quadratic convergence.

# Insufficient regularity properties in $\mathbb{H}^2$

Problem:

$$\min_{u \in \mathbb{H}^2} \mathcal{J}(u)$$

$\nabla^2 \mathcal{J}$: Lipschitz-continuous $\Rightarrow$ local quadratic convergence.

$\mathbb{H}^2$: set of square integrable processes, $\|u\|_{\mathbb{H}^2} = \sqrt{\mathbb{E}\left[\int_0^T u_t^2 \mathrm{d}t\right]}$.

- $\mathcal{J} : \mathbb{H}^2 \mapsto \mathbb{R}$ is twice differentiable.
- $\nabla \mathcal{J} : \mathbb{H}^2 \mapsto \mathbb{H}^2$.
- $\nabla^2 \mathcal{J} : \mathbb{H}^2 \mapsto \mathcal{L}(\mathbb{H}^2)$.

But $\nabla^2 \mathcal{J} : \mathbb{H}^2 \mapsto \mathcal{L}(\mathbb{H}^2)$ **not Lipschitz-continuous**.

# Counter-example

## Example

Consider $\mathcal{J}$ given by:

$$\forall u \in \mathbb{H}^2([0,1] \times \Omega, \mathbb{R}), \qquad \mathcal{J}(u) := \mathbb{E}\left[\int_0^1 l(u_t)\mathrm{d}t\right], \qquad s.t. \quad X_t^u = 0, \forall t \in [0,1],$$

where $l$ is such $l''(x) = \min(1 + |x|, 2)$; $l'(0) = 0$; $l(0) = 0$.
$\mathcal{J}$ is twice continuously differentiable, with second order-derivative $\nabla^2 \mathcal{J}$ given by $(\nabla^2 \mathcal{J}(u)(v))_t = l''(u_t)v_t$, for $u, v \in \mathbb{H}^2$.

# Counter-example

## Example

Consider $\mathcal{J}$ given by:

$$\forall u \in \mathbb{H}^2([0,1]\times\Omega,\mathbb{R}), \qquad \mathcal{J}(u) := \mathbb{E}\left[\int_0^1 l(u_t)\mathrm{d}t\right], \qquad s.t. \quad X_t^u = 0, \forall t \in [0,1],$$

where $l$ is such $l''(x) = \min(1+|x|,2)$; $l'(0)=0$; $l(0)=0$.
$\mathcal{J}$ is twice continuously differentiable, with second order-derivative $\nabla^2\mathcal{J}$ given by $(\nabla^2\mathcal{J}(u)(v))_t = l''(u_t)v_t$, for $u, v \in \mathbb{H}^2$.
However, let us define $u^{(n)} \in \mathbb{H}^2$ the constant process with value 1 with probability $1/n$ and 0 else:

$$\frac{\|\nabla^2\mathcal{J}(u^{(n)}) - \nabla^2\mathcal{J}(0)\|_{\mathcal{L}(\mathbb{H}^2)}}{\|u^{(n)}-0\|_{\mathbb{H}^2}} \geq \frac{\|\nabla^2\mathcal{J}(u^{(n)})(u^{(n)}) - \nabla^2\mathcal{J}(0)(u^{(n)})\|_{\mathbb{H}^2}}{\|u^{(n)}\|_{\mathbb{H}^2}\|u^{(n)}\|_{\mathbb{H}^2}} = \sqrt{n} \underset{n\longrightarrow+\infty}{\longrightarrow} +\infty.$$

In particular $\nabla^2\mathcal{J} : \mathbb{H}^2 \mapsto \mathcal{L}(\mathbb{H}^2)$ is not Lipschitz-continuous.

# Restriction to the space of essentially bounded processes

**As a difference with $\mathbb{R}^d$, non-equivalence of norms in our setting $\Rightarrow$ Choose right space !**

- Replace $\mathbb{H}^2$ by $\mathbb{H}^\infty$ endowed with $\|u\|_{\mathbb{H}^\infty} = \sup_{t\in[0,T]} \operatorname{essup}|u_t|$.
- Prove stability under restriction to $\mathbb{H}^\infty$:

$$\nabla\mathcal{J}(\mathbb{H}^\infty) \subset \mathbb{H}^\infty \quad ; \quad \nabla^2\mathcal{J}(\mathbb{H}^\infty) \subset \mathcal{L}(\mathbb{H}^\infty).$$

- $\mathcal{J} : \mathbb{H}^\infty \mapsto \mathbb{R}$ has **bounded and Lipschitz second-order derivative** if data regular.
- Has an impact on the backtracking line search algorithm.

## Theorem (Stability of $\mathbb{H}^\infty$)

*Under regularity, convexity and boundedness Assumptions, for all $u, v, w \in \mathbb{H}^\infty$, $X^u, Y^u, \nabla\mathcal{J}(u), \dot{X}^v, \dot{Y}^{u,v}, \nabla^2\mathcal{J}(u)(v)$ and $(\nabla^2\mathcal{J}(u))^{-1}(w)$ are in $\mathbb{H}^\infty$ and:*

$$
\begin{aligned}
&\|X^u\|_{\mathbb{H}^\infty} + \|Y^u\|_{\mathbb{H}^\infty} + \|\nabla\mathcal{J}(u)\|_{\mathbb{H}^\infty} \leq C(1 + \|u\|_{\mathbb{H}^\infty}),\\
&\|X^u - X^v\|_{\mathbb{H}^\infty} + \|Y^u - Y^v\|_{\mathbb{H}^\infty} + \|\nabla\mathcal{J}(u) - \nabla\mathcal{J}(v)\|_{\mathbb{H}^\infty} \leq C\|u - v\|_{\mathbb{H}^\infty},\\
&\|\dot{X}^v\|_{\mathbb{H}^\infty} + \|\dot{Y}^{u,v}\|_{\mathbb{H}^\infty} + \|\nabla^2\mathcal{J}(u)(v)\|_{\mathbb{H}^\infty} \leq C\|v\|_{\mathbb{H}^\infty},\\
&\|\dot{Y}^{u,w} - \dot{Y}^{v,w}\|_{\mathbb{H}^\infty} + \|\nabla^2\mathcal{J}(u)(w) - \nabla^2\mathcal{J}(v)(w)\|_{\mathbb{H}^\infty} \leq C\|u - v\|_{\mathbb{H}^\infty}\|w\|_{\mathbb{H}^\infty},\\
&\|(\nabla^2\mathcal{J}(u))^{-1}(w)\|_{\mathbb{H}^\infty} \leq C\|w\|_{\mathbb{H}^\infty}.
\end{aligned}
$$

- *$\nabla^2\mathcal{J}$ defines a Lipschitz-continuous operator from $\mathbb{H}^\infty \mapsto \mathcal{L}(\mathbb{H}^\infty)$.*
- *$\nabla^2\mathcal{J}(u)$ and $(\nabla^2\mathcal{J}(u))^{-1}$ are bounded linear operators, uniformly in $u$.*
- *The Newton direction $\Delta_u$ at the point $u \in \mathbb{H}^\infty$ is in $\mathbb{H}^\infty$.*

# Backtracking line search

**Alg. 1: Standard Backtracking line search**

1: **Inputs:** Current point $u \in \mathbb{H}^\infty$, Current search direction $\Delta_u \in \mathbb{H}^\infty$, $\beta \in (0,1)$, $\gamma \in (0,1)$.
2: $\sigma = 1$.
3: **while** $\mathcal{J}(u + \sigma\Delta_u) > \mathcal{J}(u) + \gamma\sigma\langle\nabla\mathcal{J}(u), \Delta_u\rangle_{\mathbb{H}^2}$ **do**
4: $\quad\sigma \leftarrow \beta\sigma$.
5: **end while**
6: **return** $u + \sigma\Delta_u$.

The global convergence of the method is not guaranteed in our setting. Open problem

## Alg 2: Gradient Backtracking line search

1: **Inputs:** Current point $u \in \mathbb{H}^\infty$, Current search direction $\Delta_u \in \mathbb{H}^\infty$, $\beta \in (0,1)$, $\gamma \in (0,1)$.
2: $\sigma = 1$.
3: **while** $\|\nabla \mathcal{J}(u + \sigma\Delta_u)\|_{\mathbb{H}^\infty} > (1 - \gamma\sigma)\|\nabla \mathcal{J}(u)\|_{\mathbb{H}^\infty}$ **do**
4: $\quad \sigma \leftarrow \beta\sigma$.
5: **end while**
6: **return** $u + \sigma\Delta_u$.

## Alg 2: Gradient Backtracking line search

1: **Inputs:** Current point $u \in \mathbb{H}^\infty$, Current search direction $\Delta_u \in \mathbb{H}^\infty$, $\beta \in (0,1)$, $\gamma \in (0,1)$.
2: $\sigma = 1$.
3: **while** $\|\nabla \mathcal{J}(u + \sigma\Delta_u)\|_{\mathbb{H}^\infty} > (1 - \gamma\sigma)\|\nabla \mathcal{J}(u)\|_{\mathbb{H}^\infty}$ **do**
4: $\quad \sigma \leftarrow \beta\sigma$.
5: **end while**
6: **return** $u + \sigma\Delta_u$.

### Theorem (global convergence and locally quadratic)

▷ *Alg. 2 terminates in finitely many iterations.*
▷ *If Alg. 2 returns $\sigma = 1$, then the new point $u + \Delta_u$ satisfies:*

$$\|\nabla \mathcal{J}(u + \Delta_u)\|_{\mathbb{H}^\infty} \leq \min(1 - \gamma, C\|\nabla \mathcal{J}(u)\|_{\mathbb{H}^\infty})\|\nabla \mathcal{J}(u)\|_{\mathbb{H}^\infty}.$$

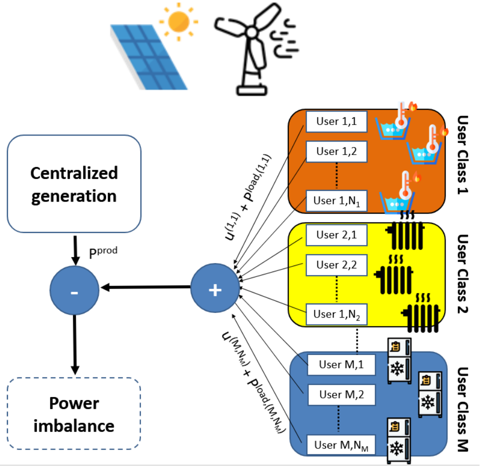
▷ *If Alg. 2 returns $\sigma < 1$, then the new iterate $u + \sigma\Delta_u$ satisfies:*

$$\|\nabla \mathcal{J}(u + \sigma\Delta_u)\|_{\mathbb{H}^\infty} \leq \|\nabla \mathcal{J}(u)\|_{\mathbb{H}^\infty} - \frac{\beta\gamma(1-\gamma)}{C}.$$

*$C$ is a constant depending on data.*

1. Recall on Newton method for optimization in $\mathbb{R}^d$
2. Stochastic control: framework and Newton step computation
3. Theoretical convergence properties
4. Numerical implementation and results

# Numerical example: Set of heterogeneous consumers, with different flexibility - Supply-demand balance




- Common noise (weather...)
- Independent individual noise (agent consumption...)
- General filtration (not necessarily Brownian) to allow jumps in exogenous factors

# Control/dynamics: multi-category with common noise

- $M \in \mathbb{N}$: number of agent categories, indexed by $k$ or $l$
- Each category $k \in \{1, ..., M\}$ (with same characteristics) has $N_k$ agents indexed by $i$ or $j$

# Control/dynamics: multi-category with common noise

- $M \in \mathbb{N}$: number of agent categories, indexed by $k$ or $l$
- Each category $k \in \{1, ..., M\}$ (with same characteristics) has $N_k$ agents indexed by $i$ or $j$
- Dynamics for storage $i$ in category $k$

$$X_t^{(k,i)} = x_0^{(k,i)} + \int_0^t \left( \alpha_s^{(k)} u_s^{(k,i)} + \beta_s^{(k)} X_s^{(k,i)} + \gamma_s^{(k,i)}(\omega) \right) \mathrm{d}s,$$

- Controls : $(u^{(k,i)})_{1 \leq k \leq M, 1 \leq i \leq N_k}$ power consumptions

# Control/dynamics: multi-category with common noise

- $M \in \mathbb{N}$: number of agent categories, indexed by $k$ or $l$
- Each category $k \in \{1, ..., M\}$ (with same characteristics) has $N_k$ agents indexed by $i$ or $j$
- Dynamics for storage $i$ in category $k$

$$X_t^{(k,i)} = x_0^{(k,i)} + \int_0^t \left( \alpha_s^{(k)} u_s^{(k,i)} + \beta_s^{(k)} X_s^{(k,i)} + \gamma_s^{(k,i)}(\omega) \right) \mathrm{d}s,$$

- Controls : $(u^{(k,i)})_{1 \le k \le M, 1 \le i \le N_k}$ power consumptions
- Examples of storage:
  - Battery state of charge: $X_t = x_0 + \int_0^t \frac{u_s}{\mathcal{E}_{max}} \mathrm{d}s$.
  - Temperature thermal storage: $X_t = x_0 + \int_0^t (\alpha u_s - \beta(X_s - T_{out}(s)) + \gamma_s) \mathrm{d}s$.
- Processes impacting individual consumers (consumption, solar production...) are independent given $\mathcal{G}_T$ (weather noise. . . )

# Objective function (to be minimised): a **cooperative** approach

$$\mathbb{E}\left[\frac{1}{N}\sum_{k=1}^{M}\sum_{i=1}^{N_k}\underbrace{\left\{\int_0^T Q^{(k,i)}\left(t, u_t^{(k,i)}, X_t^{(k,i)}\right)\mathrm{d}t + \Psi^{(k,i)}\left(X_T^{(k,i)}\right)\right\}}_{\text{management cost for storage } i \text{ of category } k}\right]$$

$$+\,\mathbb{E}\left[\int_0^T \underbrace{\mathcal{L}\left(t, \frac{1}{N}\sum_{l=1}^{M}\sum_{j=1}^{N_l}\left(u_t^{(l,j)} + \mathrm{P}_t^{\mathtt{load},(l,j)}\right) - \mathrm{P}_t^{\mathtt{prod}}\right)}_{\text{instantaneous overall imbalance}}\mathrm{d}t\right],$$

$$Q^{(k,i)}(t,u,x) = \mu_t^{(k)}\left(u - u_t^{\mathtt{ref},(k,i)}\right)^2/2 + \nu_t^{(k)}\left(x - x_t^{\mathtt{ref},(k,i)}\right)^2/2,$$

$$\Psi^{(k,i)}(x) = \rho^{(k)}\left(x - x_T^{\mathtt{f},(k,i)}\right)^2/2.$$

# The control problem

$$\min_{(u^{(k,i)})_{k,i}} \mathbb{E}\left[\frac{1}{N}\sum_{k=1}^{M}\sum_{i=1}^{N_k}\left\{\int_0^T Q^{(k,i)}\left(t, u_t^{(k,i)}, X_t^{(k,i)}\right)\mathrm{d}t + \Psi^{(k,i)}\left(X_T^{(k,i)}\right)\right\}\right]$$
$$+\mathbb{E}\left[\int_0^T \mathcal{L}\left(t, \frac{1}{N}\sum_{l=1}^{M}\sum_{j=1}^{N_l}(u_t^{(l,j)} + \mathrm{P}_t^{\mathtt{load},(l,j)}) - \mathrm{P}_t^{\mathtt{prod}}\right)\mathrm{d}t\right],$$
$$s.t. \quad X_t^{(k,i)} = x_0^{(k,i)} + \int_0^t \left(\alpha_s^{(k)} u_s^{(k,i)} + \beta_s^{(k)} X_s^{(k,i)} + \gamma_s^{(k,i)}\right)\mathrm{d}s, \ \forall k, i.$$

- Semi Linear-Quadratic ($\mathcal{L}(t,.)$ not necessarily quadratic) and strongly convex stochastic control problem...
- ... in high dimension ($N := \sum_{k=1}^{M} N_k$) with coupling.

# Approximation for the aggregator and consumers

From EG, MG: "Federated stochastic control of numerous heterogeneous energy storage systems" https://hal.archives-ouvertes.fr/hal-03108611

- Theorem: the $N$-dimensional control problem is approximately equivalent to a leader-follower control problem:
  - 1 control problem for the aggregator in dimension $M$ (number of categories)
  - for each consumer, a 1-dimensional control problem

Aggregator gives a coordination signal, to be used by all consumers in parallel $\Rightarrow$ Solves the privacy preserving issue

# Approximation for the aggregator and consumers

From EG, MG: "Federated stochastic control of numerous heterogeneous energy storage systems" https://hal.archives-ouvertes.fr/hal-03108611

- Theorem: the $N$-dimensional control problem is approximately equivalent to a leader-follower control problem:
  - 1 control problem for the aggregator in dimension $M$ (number of categories)
  - for each consumer, a 1-dimensional control problem

  Aggregator gives a coordination signal, to be used by all consumers in parallel $\Rightarrow$ Solves the privacy preserving issue

- Theorem: error bounds
  - $\mathcal{O}(1/\sqrt{M})$ for control, in $L_2$
  - $\mathcal{O}(1/M)$ for functional cost

# The aggregator control problem

$$\min_{(u^{(k)})_{1\le k\le M}} \mathbb{E}\left[\sum_{k=1}^{M}\pi^{(k)}\left\{\int_0^T \bar{Q}^{(k,N)}\left(t, u_t^{(k)}, X_t^{(k)}\right)\mathrm{d}t + \bar{\Psi}^{(k,N)}\left(X_T^{(k)}\right)\right\}\right]$$

$$+\mathbb{E}\left[\int_0^T \mathcal{L}\left(t, \underbrace{\sum_{l=1}^{M}\pi^{(l)}u_t^{(l)} + \bar{\mathrm{P}}_t^{\mathtt{load},(N)} - \mathrm{P}_t^{\mathtt{prod}}}_{:=\bar{v}^{(N)},\ \text{coordination signal}}\right)\mathrm{d}t\right],$$

$$s.t. \quad X_t^{(k)} = \bar{x}_0^{(k,N)} + \int_0^t \left(\alpha_s^{(k)}u_s^{(k)} + \beta_s^{(k)}X_s^{(k)} + \bar{\gamma}_s^{(k,N)}\right)\mathrm{d}s,\ \forall k,$$

$$\bar{Q}^{(k,N)}(t,u,x) := \frac{1}{2}\left(\mu_t^{(k)}\left(u - \bar{u}_t^{\mathtt{ref},(k,N)}\right)^2 + \nu_t^{(k)}\left(x - \bar{x}_t^{\mathtt{ref},(k,N)}\right)^2\right),$$

$$\bar{\Psi}^{(k,N)}(x) := \rho^{(k)}\left(x - \bar{x}_T^{\mathtt{f},(k,N)}\right)^2/2.$$

Can be solved in the $\mathbb{G}$-filtration. Depends on statistics of consumers. Reduced-size problem.

# Solving the aggregator problem

**Description of Newton iteration:**

- Each iteration requires solving one ODE and three BSDEs.

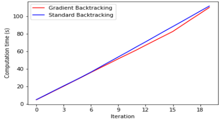
- **Numerical solution of BSDEs**. Many conditional expectations $\rightsquigarrow$ Empirical Least-Squares Regression [8].
  **Markovian framework:**

$$Y_t = \mathbb{E}\left[\int_t^T f(s, X_s, Y_s)\mathrm{d}s + G(X_T)|\mathcal{F}_t\right] = \phi_t(X_t),$$

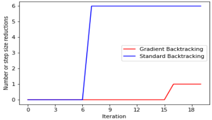
$$\phi_t(.) = \arg\min_{h \text{ meas.}} \mathbb{E}\left[\left(\int_t^T f(s, X_s, Y_s)\mathrm{d}s + G(X_T) - h(X_t)\right)^2\right].$$

  - Solve backwards in $t$ (after time discretization $\simeq$ Euler).
  - Choose $h$ in finite-dimensional functional vector space $V$ or non linear space (e.g. NN).
  - Expectation $\simeq$ empirical mean over the simulations
  - $\mathbb{H}^\infty$ norm computed over the simulations

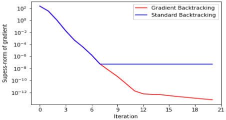
# Numerical performance



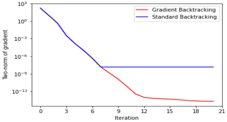

(a) Computation time (in seconds)



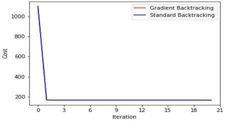

(b) Number of step size reductions



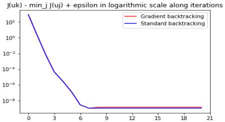

(c) $\|\nabla \mathcal{J}(u^{(k)})\|_{\mathbb{H}^\infty}$ along iterations




(d) $\|\nabla \mathcal{J}(u^{(k)})\|_{\mathbb{H}^2}$ along iterations




(e) Cost $\bar{\mathcal{J}}(u^{(k)})$ along iterations $k$




(f) Suboptimality gap $\bar{\mathcal{J}}(u^{(k)}) - \min_j \bar{\mathcal{J}}(u^{(j)}) + 10^{-9}$ (log scale)

Figure: Performance of Newton method with the two line search methods along iterations

# Conclusion and perspectives

- Design of Newton algorithm for stochastic control
- Careful choice of norms
- Proof of convergence (global, and locally quadratic). A few iterations are sufficient.
- Iterative method made of simple BSDEs, fast to solve.
- On-going works:
  - extension to more general control problems
  - analysis of numerical errors

# References I

C. Bender, J. Zhang, et al., *Time discretization and Markovian iteration for coupled FBSDEs*, The Annals of Applied Probability, 18 (2008), pp. 143–177.

J.-M. Bismut, *Linear quadratic optimal stochastic control with random coefficients*, SIAM Journal on Control and Optimization, 14 (1976), pp. 419–444.

S. Boyd and L. Vandenberghe, *Convex optimization*, Cambridge university press, 2004.

H. Brezis, *Functional analysis, Sobolev spaces and partial differential equations*, Springer Science+Business Media, 2010.

F. Chernousko and A. Lyubushin, *Method of successive approximations for solution of optimal control problems*, Optimal Control Applications and Methods, 3 (1982), pp. 101–114.

E. Gobet and M. Grangereau, *Extended McKean-Vlasov optimal stochastic control applied to smart grid management*, to appear in ESAIM: Control, Optimisation and Calculus of Variations, (2021).

# References II

E. Gobet and M. Grangereau, *Federated stochastic control of numerous heterogeneous energy storage systems*, preprint hal-03108611, (2021).

E. Gobet and P. Turkedjiev, *Linear regression MDP scheme for discrete backward stochastic differential equations under general conditions*, Mathematics of Computation, 85 (2016), pp. 1359–1391.

J. Han, A. Jentzen, and E. Weinan, *Solving high-dimensional partial differential equations using deep learning*, Proceedings of the National Academy of Sciences, 115 (2018), pp. 8505–8510.

S. Ji, S. Peng, Y. Peng, and X. Zhang, *Three algorithms for solving high-dimensional fully-coupled FBSDEs through deep learning*, IEEE Intelligent Systems, (2020).

B. Kerimkulov, D. Šiška, and Ł. Szpruch, *A modified MSA for stochastic control problems*, arXiv preprint arXiv:2007.05209, (2020).

J. Nocedal and S. Wright, *Numerical optimization*, Springer Science & Business Media, 2006.

# References III

S. Peng and Z. Wu, *Fully coupled forward-backward stochastic differential equations and applications to optimal control*, SIAM Journal on Control and Optimization, 37 (1999), pp. 825–843.

J. Yong, *Linear forward-backward stochastic differential equations with random coefficients*, Probability theory and related fields, 135 (2006), pp. 53–83.

J. Zhang, *Backward stochastic differential equations*, in Backward Stochastic Differential Equations, Springer, 2017, pp. 79–99.