Prediction and Second Order PDEs

Erhan Bayraktar ¹
(with I. Ekren ¹, X. Zhang ², and Y. Zhang ¹)

¹University of Michigan
²University of Vienna

June 2023
Stochastic control and Financial Engineering
Princeton
Outline

1 Introduction

2 Full Information
   - The Geometric stopping problem
   - Long time asymptotics
   - Our Main contribution
   - Finite vs. geometric regret conjecture

3 Partial Information Set-up
   - Derivatives in Wasserstein space
   - Limiting PDE
   - Verification approach
   - Viscosity approach
Outline

1. Introduction

2. Full Information
   - The Geometric stopping problem
   - Long time asymptotics
   - Our Main contribution
   - Finite vs. geometric regret conjecture

3. Partial Information Set-up
   - Derivatives in Wasserstein space
   - Limiting PDE
   - Verification approach
   - Viscosity approach
Introduction: Prediction with expert advice

A sequential decision problem in an adversarial set-up.

At each time step $t \in \{1, \ldots, T\}$

- A forecaster chooses an expert $I_t \in \{1, \ldots, K\}$ from a set of $K$ experts.
- An adversary (nature) simultaneously chooses the set $J_t \subset \{1, \ldots, K\}$ of winning experts meaning

  $$g^i_t = 1 \text{ if } i \in J_t \text{ and } g^i_t = 0 \text{ if } i \notin J_t.$$ 

- The gain of the forecaster is $g^i_{I_t}.$
- The forecaster’s goal is to minimize his regret:

  $$R_T = \max_{i=1,\ldots,K} G^i_T - G_T = \max_{i=1,\ldots,K} \sum_{t=1}^{T} (g^i_t - g^i_{I_t})$$

  where $G^i_T$ is the total gain of the expert $i$ and $G_T$ is the gain of the forecaster at the final time $T.$

- The objective of the adversary is to choose the winning experts to maximize the regret of the forecaster.
Prediction with expert advice

A sequential decision problem in an adversarial set-up.

At each time step $t \in \{1, \ldots, T\}$

- A forecaster chooses an expert $I_t \in \{1, \ldots, K\}$ from a set of $K$ experts.
- An adversary (nature) simultaneously chooses the set $J_t \subset \{1, \ldots, K\}$ of winning experts meaning
  \[
g^i_t = 1 \text{ if } i \in J_t \text{ and } g^i_t = 0 \text{ if } i \notin J_t.
  \]

- The gain of the forecaster is $g^I_{lt}$.
- The forecaster’s goal is to minimize his regret:
  \[
  R_T = \max_{i=1,\ldots,K} G^i_T - G_T = \max_{i=1,\ldots,K} \sum_{t=1}^T (g^i_t - g^I_{lt})
  \]
  where $G^i_T$ is the total gain of the expert $i$ and $G_T$ is the gain of the forecaster at the final time $T$.

- The objective of the adversary is to choose the winning experts to maximize the regret of the forecaster.
Introduction: Prediction with expert advice
A sequential decision problem in an adversarial set-up.

At each time step \( t \in \{1, \ldots, T\} \)

- A forecaster chooses an expert \( I_t \in \{1, \ldots, K\} \) from a set of \( K \) experts.
- An adversary (nature) simultaneously chooses the set \( J_t \subset \{1, \ldots, K\} \) of winning experts meaning
  \[
  g_t^i = 1 \text{ if } i \in J_t \text{ and } g_t^i = 0 \text{ if } i \notin J_t.
  \]
- The gain of the forecaster is \( g_{lt}^i \).
- The forecaster’s goal is to minimize his regret:
  \[
  R_T = \max_{i=1,\ldots,K} G_T^i - G_T = \max_{i=1,\ldots,K} \sum_{t=1}^{T} (g_t^i - g_{lt}^i)
  \]
  where \( G_T^i \) is the total gain of the expert \( i \) and \( G_T \) is the gain of the forecaster at the final time \( T \).
- The objective of the adversary is to choose the winning experts to maximize the regret of the forecaster.
Introduction: Prediction with expert advice

A sequential decision problem in an adversarial set-up.

At each time step $t \in \{1, \ldots, T\}$

- A forecaster chooses an expert $I_t \in \{1, \ldots, K\}$ from a set of $K$ experts.
- An adversary (nature) simultaneously chooses the set $J_t \subset \{1, \ldots, K\}$ of winning experts meaning
  \[
  g^i_t = 1 \text{ if } i \in J_t \text{ and } g^i_t = 0 \text{ if } i \notin J_t.
  \]
- The gain of the forecaster is $g^{I_t}_t$.
- The forecaster’s goal is to minimize his regret:
  \[
  R_T = \max_{i=1,\ldots,K} G^i_T - G_T = \max_{i=1,\ldots,K} \sum_{t=1}^T (g^i_t - g^{I_t}_t)
  \]
  where $G^i_T$ is the total gain of the expert $i$ and $G_T$ is the gain of the forecaster at the final time $T$.
- The objective of the adversary is to choose the winning experts to maximize the regret of the forecaster.
Introduction: Prediction with expert advice
A sequential decision problem in an adversarial set-up.

At each time step $t \in \{1, \ldots, T\}$

- A forecaster chooses an expert $l_t \in \{1, \ldots, K\}$ from a set of $K$ experts.
- An adversary (nature) simultaneously chooses the set $J_t \subset \{1, \ldots, K\}$ of winning experts meaning
  \[ g^i_t = 1 \text{ if } i \in J_t \text{ and } g^i_t = 0 \text{ if } i \notin J_t. \]
- The gain of the forecaster is $g^l_t$.
- The forecaster’s goal is to minimize his regret:
  \[ R_T = \max_{i=1,\ldots,K} G^i_T - G_T = \max_{i=1,\ldots,K} \sum_{t=1}^{T} (g^i_t - g^l_t) \]
  where $G^i_T$ is the total gain of the expert $i$ and $G_T$ is the gain of the forecaster at the final time $T$.
- The objective of the adversary is to choose the winning experts to maximize the regret of the forecaster.
Introduction: Prediction with expert advice

- In classical stochastic games, statistical assumptions are made on the realizations of $g^i_t$ and the agents optimize some expected “return”.
- In the context of online learning, we do not make any statistical assumptions on $g^i_t$. The criterion $R_T$ is ”how much the forecaster regrets, in hindsight, of not having followed the advice of the best expert”.
- $R_T$ is commonly used as a criterion expressing the gap of optimality (with respect to the best constant strategy) in online (dynamic) allocation problems with learning.
- The forecaster chooses $I_t$ at time $t$ to obtain
  \[ G_T = \sum_{t=1}^{T} g^I_t. \]
- But learns the final benchmark at time $T$
  \[ \max_{i=1,\ldots,K} \sum_{t=1}^{T} g^i_t. \]
Introduction: Prediction with expert advice

- In classical stochastic games, statistical assumptions are made on the realizations of $g_t^i$ and the agents optimize some expected "return".
- In the context of online learning, we do not make any statistical assumptions on $g_t^i$. The criterion $R_T$ is "how much the forecaster regrets, in hindsight, of not having followed the advice of the best expert".
- $R_T$ is commonly used as a criterion expressing the gap of optimality (with respect to the best constant strategy) in online (dynamic) allocation problems with learning.
- The forecaster chooses $I_t$ at time $t$ to obtain

$$G_T = \sum_{t=1}^T g_t^{I_t}.$$ 

- But learns the final benchmark at time $T$

$$\max_{i=1,...,K} \sum_{t=1}^T g_t^i.$$
Two different games for prediction with expert advice

- Full information setting.
  - At each time step, the forecaster learns how much each expert is winning.
  - $X_t$ where $X_t^i = G_t^i - G_t$ (gap of optimality with respect to each expert) is the relevant state variable to state the problem as a stochastic game where the strategy of the adversary is $\alpha$, the strategy of the forecaster is $\beta$, and the criterion is

$$\mathbb{E}^{\alpha, \beta} [\Phi(X_T)]$$

where $\Phi(x) = \max_{i=1,\ldots,K} x^i$.

- Partial information set-up: similar to Adversarial (multi-armed) bandit setting.
  - The forecaster only learns if the expert he chose has won (no information on other experts!)
  - The relevant state variable is the distribution $m_t$ of $X_t$ conditional on the information of the forecaster and the criterion is

$$\mathbb{E}^{\alpha, \beta} [\Phi(X_T)] = \mathbb{E}^{\alpha, \beta} \left[ \int \Phi(x) m_T(dx) \right] .$$
Questions

- What is the optimal algorithm and the regret value?
- What is the growth of the regret in $T$?
- Does the optimal algorithm have a succinct and intuitive description (even for 2 experts)?
- What are the hardest (adversarial) sequences of experts gains and do they follow a succinct pattern?
[Cover, 1966] completely solves the problem for \( K = 2 \).

Adversary chooses \((g^1_t = 1, g^2_t = 0)\) with probability \( \frac{1}{2} \) and \((g^1_t = 0, g^2_t = 1)\) with probability \( \frac{1}{2} \).

Optimal algorithm for the forecaster is to choose the leading expert and the lagging expert with probabilities depending on \( T \), the gap of optimality \( G^{(1)}_t - G^{(2)}_t \), and \( t \).

The growth of regret is \( \sqrt{\frac{T}{2\pi}} \).
The famous **multiplicative weights** algorithm which chooses expert $i$ with probability

$$
\frac{e^{\eta G^i_t}}{\sum_j e^{\eta G^j_t}}
$$

is a standard algorithm for general $K$ (see [Cesa-Bianchi et al., 1997], [Cesa-Bianchi and Lugosi, 2006])

- As $(K, T) \to \infty$ the asymptotic minimax optimal regret rate is $\sqrt{(T/2)\log K}$.
- The asymptotic bound is **attained** by the MWA (**multiplicative weights** algorithm).
For finite \( K \), the MWA underperforms significantly.

For \( K = 2 \) the regret of a MWA learner is 47% larger than the optimal learner [Gravin et al., 2017])

[Gravin et al., 2016] finally to solve this problem for \( K = 3 \) (the regret of the MWA learner is 39% larger than the optimal.)

For \( K = 3 \), they show that the optimal adversary is the so-called COMB strategy and conjectured that it is also optimal for \( K \geq 4 \).

Using PDE and stochastic control tools, we solve the open problem for \( K = 4 \) as \( T \to \infty \) (the regret of the MWA learner is 25% larger than the optimal.)

- B., Ekren, and Y. Zhang [AAP, 2020]: geometric horizon.
- B., Ekren, and X. Zhang [Communication in PDEs, 2020]: finite horizon.

Extensions.

- B., Ekren, and X. Zhang [JMLR, 2021]: the adversary can corrupt the distribution of outcomes.
- B., Poor, and X. Zhang [IEEE TIT, 2020]: Multiplicative expert vs a malicious expert.
For finite $K$, the MWA underperforms significantly.

For $K=2$ the regret of a MWA learner is 47% larger than the optimal learner [Gravin et al., 2017])

[Gravin et al., 2016] finally to solve this problem for $K = 3$ (the regret of the MWA learner is 39% larger than the optimal.)

For $K = 3$, they show that the optimal adversary is the so-called COMB strategy and conjectured that it is also optimal for $K \geq 4$.

Using PDE and stochastic control tools, we solve the open problem for $K = 4$ as $T \to \infty$ (the regret of the MWA learner is 25% larger than the optimal.)

▶ B., Ekren, and Y. Zhang [AAP, 2020]: geometric horizon.
▶ B., Ekren, and X. Zhang [Communication in PDEs, 2020]: finite horizon.

Extensions.

▶ B., Ekren, and X. Zhang [JMLR, 2021] : the adversary can corrupt the distribution of outcomes.
Outline

1 Introduction

2 Full Information
   - The Geometric stopping problem
   - Long time asymptotics
   - Our Main contribution
   - Finite vs. geometric regret conjecture

3 Partial Information Set-up
   - Derivatives in Wasserstein space
   - Limiting PDE
   - Verification approach
   - Viscosity approach
Set Up

- Both parties observe the same source of information \( \{(G^i_s, G_s) : s \leq t\} \) and their decisions are revealed simultaneously.
- They choose their control for the next step.
- When the adversary chooses \( J_t \subset \{1, \ldots, K\} \), the gain of each expert in \( J_t \) increases by 1 i.e.,

\[
\begin{align*}
G^i_t &= G^i_{t-1} + 1 \quad \text{if} \ i \in J_t \\
G^i_t &= G^i_{t-1} \quad \text{if} \ i \notin J_t.
\end{align*}
\]

- When the forecaster chooses \( I_t \in \{1, \ldots, K\} \), the gain of forecaster also increases if \( I_t \in J_t \) i.e.,

\[
\begin{align*}
G_t &= G_{t-1} + 1 \quad \text{if} \ I_t \in J_t \\
G_t &= G_{t-1} \quad \text{if} \ I_t \notin J_t.
\end{align*}
\]

- The relevant state variable is \( X^i_t = G^i_t - G_t, \ i = 1, \ldots, K \).
Mixed strategies

- All controls are mixed strategies. Given the information the adversary chooses a distribution on \(2\{1, \ldots, K\}\) denote by \(\alpha_t\), the forecaster chooses a distribution on \(\{1, \ldots, K\}\), denote by \(\beta_t\), and repeat the game \(T\) times.

- The problem is

\[
\inf_{\beta} \sup_{\alpha} E^{\alpha, \beta}\left[ \max_{i=1, \ldots, K} G^i_T - G_T \right] = \sup_{\alpha} \inf_{\beta} E^{\alpha, \beta}\left[ \max_{i=1, \ldots, K} G^i_T - G_T \right].
\]

- Starting from any state \(X_0 = x \in \mathbb{R}^K\), we can define the value function

\[
V(T, X_0) = \min_{\text{forecaster}} \max_{\text{adversary}} E^{\alpha, \beta}\left[ \Phi(X_T) \right] = \max_{\text{adversary}} \min_{\text{forecaster}} E^{\alpha, \beta}\left[ \Phi(X_T) \right].
\]
Mixed strategies

- All controls are mixed strategies. Given the information the adversary chooses a distribution on \( 2\{1, \ldots, K\} \) denote by \( \alpha_t \), the forecaster chooses a distribution on \( \{1, \ldots, K\} \), denote by \( \beta_t \), and repeat the game \( T \) times.

- The problem is

\[
\inf_{\beta} \sup_{\alpha} \mathbb{E}^{\alpha, \beta} \left[ \max_{i=1, \ldots, K} G^i_T - G_T \right] = \sup_{\alpha} \inf_{\beta} \mathbb{E}^{\alpha, \beta} \left[ \max_{i=1, \ldots, K} G^i_T - G_T \right].
\]

- Starting from any state \( X_0 = x \in \mathbb{R}^K \), we can define the value function

\[
V(T, X_0) = \min_{\text{forecaster}} \max_{\text{adversary}} \mathbb{E}^{\alpha, \beta} [\Phi(X_T)]
\]

\[
= \max_{\text{adversary}} \min_{\text{forecaster}} \mathbb{E}^{\alpha, \beta} [\Phi(X_T)].
\]
Mixed strategies

- All controls are mixed strategies. Given the information the adversary chooses a distribution on $2\{1,\ldots,K\}$ denote by $\alpha_t$, the forecaster chooses a distribution on $\{1,\ldots,K\}$, denote by $\beta_t$, and repeat the game $T$ times.

- The problem is

$$
\inf_{\beta} \sup_{\alpha} \mathbb{E}^{\alpha,\beta} \left[ \max_{i=1,\ldots,K} G^i_T - G_T \right] = \sup_{\alpha} \inf_{\beta} \mathbb{E}^{\alpha,\beta} \left[ \max_{i=1,\ldots,K} G^i_T - G_T \right].
$$

- Starting from any state $X_0 = x \in \mathbb{R}^K$, we can define the value function

$$
V(T, X_0) = \min_{\text{forecaster}} \max_{\text{adversary}} \mathbb{E}^{\alpha,\beta} [\Phi(X_T)]
$$

$$
= \max_{\text{adversary}} \min_{\text{forecaster}} \mathbb{E}^{\alpha,\beta} [\Phi(X_T)].
$$
The Geometric stopping

- We consider the problem where now $\tau$ is a geometric random variable with parameter $\delta \to 0$

$$V^\delta(x) = \min_{\text{forecaster}} \max_{\text{adversary}} \mathbb{E}^{\alpha,\beta} [\Phi(X_\tau)]$$

that satisfies the DPP

$$V^\delta(x) = \delta \Phi(x) + (1 - \delta) \inf_{\beta \in \mathcal{B}} \sup_{\alpha \in \mathcal{A}} \sum_{J \subset \{1,\ldots,K\}} \alpha_J \left( V^\delta(x + e_J) - \beta(J) \right).$$

- Here $e_J = \sum_{j \in J} e_j$, where $(e_1, \ldots, e_K)$ is the standard basis of $\mathbb{R}^K$. 
The optimal strategies for $K = 2$ and $3$

- [Cover, 1966] describes the optimal strategies for both the forecaster and the adversary for $K = 2$.
- The adversary chooses each expert with probability $\frac{1}{2}$.
- **Probability Matching Algorithm**: The forecaster chooses the expert $i$ with probability

$$\mathbb{P}(\text{expert } i \text{ is the winning expert at the end of the game})$$

- [Gravin et al., 2016] describes the strategies for both the forecaster and adversary by computing $V^\delta$ based on a guess and verify for $K = 3$.
- The adversary chooses the COMB strategy

$$J = \{\text{leading expert, third leading expert,...}\} \text{ with probability } \frac{1}{2} \text{ and } J^c = \{\text{second leading expert, forth leading expert,...}\} \text{ with probability } \frac{1}{2}$$
The optimal strategies for $K = 2$ and $3$

- [Cover, 1966] describes the optimal strategies for both the forecaster and the adversary for $K = 2$.
- The adversary chooses each expert with probability $\frac{1}{2}$.
- **Probability Matching Algorithm**: The forecaster chooses the expert $i$ with probability

  \[ P(\text{expert } i \text{ is the winning expert at the end of the game}) \]

- [Gravin et al., 2016] describes the strategies for both the forecaster and adversary by computing $V^\delta$ based on a guess and verify for $K = 3$.
  - The adversary chooses the COMB strategy

\[ J = \{\text{leading expert, third leading expert,} \ldots \} \text{ with probability } \frac{1}{2} \text{ and} \]
\[ J^c = \{\text{second leading expert, forth leading expert,} \ldots \} \text{ with probability } \frac{1}{2} \]
The optimal strategies for $K = 2$ and $3$

- [Cover, 1966] describes the optimal strategies for both the forecaster and the adversary for $K = 2$.
- The adversary chooses each expert with probability $\frac{1}{2}$.
- **Probability Matching Algorithm:** The forecaster chooses the expert $i$ with probability

  $$P(\text{expert } i \text{ is the winning expert at the end of the game})$$

- [Gravin et al., 2016] describes the strategies for both the forecaster and adversary by computing $V^\delta$ based on a guess and verify for $K = 3$.
- The adversary chooses the **COMB** strategy

  $$J = \{\text{leading expert, third leading expert,} \ldots\} \text{ with probability } \frac{1}{2} \text{ and }$$
  $$J^c = \{\text{second leading expert, forth leading expert,} \ldots\} \text{ with probability } \frac{1}{2}$$
Long time asymptotics ($\delta \to 0$)

- We define the rescaled value function $u^\delta(x) = V^\delta \left( \frac{x}{\sqrt{\delta}} \right) \sqrt{\delta}$.
- After some algebra the DPP becomes

$$u^\delta(x) - \Phi(x) = (1 - \delta) \inf_{\beta \in \mathcal{B}} \sup_{\alpha \in \mathcal{A}} \sum_{J \subset \{1, \ldots, K\}} \alpha_J \frac{u^\delta \left( x + \sqrt{\delta} e_J \right) - u^\delta(x) - \sqrt{\delta} \beta(J)}{\delta}.$$

- Assuming the convergence happens,

$$u^\delta \left( x + \sqrt{\delta} e_J \right) - u^\delta(x) - \sqrt{\delta} \beta(J) \sim \frac{1}{2} e_J^\top D^2_{xx} u^\delta e_J + \frac{1}{\sqrt{\delta}} \left( e_J^\top D_x u^\delta(x) - \beta(J) \right)$$

and the DPP becomes

$$u^\delta(x) - \Phi(x) = (1 - \delta) \inf_{\beta \in \mathcal{B}} \sup_{\alpha \in \mathcal{A}} \sum_{J \subset \{1, \ldots, K\}} \alpha_J \left( \frac{1}{2} e_J^\top D^2_{xx} u^\delta e_J + \frac{1}{\sqrt{\delta}} \left( e_J^\top D_x u^\delta(x) - \beta(J) \right) \right).$$
Properties of the limiting problem

- In order to cancel the effect of the term $\frac{1}{\sqrt{\delta}}(e^T J \partial_x u(x) - \beta(J))$ where $u(x) = \lim_{\delta \downarrow 0} u^\delta(x)$, forecaster has to choose $\beta_i = \partial_{x_i} u(x)$ at point $x \in \mathbb{R}^K$.

- The adversary chooses the winning experts by finding the unique viscosity solution the elliptic PDE

$$u(x) - \Phi(x) = \sup_{J \subset \{1, \ldots, K\}} \sum_{\alpha \in \mathcal{A}} \alpha_J \left( \frac{1}{2} e_J^T \partial_{xx} u e_J \right)$$

that has linear growth.

- The limiting problem is fully described by this PDE

$$u(x) - \sup_{J \subset \{1, \ldots, K\}} \frac{1}{2} e_J^T \partial_{xx} u e_J = \Phi(x).$$

- Convergence of a numerical scheme as the Barles-Souganidis method.
The advantage of the long time asymptotics

- For $\delta > 0$, the problem in hand is a game between the forecaster and the adversary.
- As $\delta \downarrow 0$ we only have a stochastic control problem for the adversary. The forecaster has to choose $\beta$ given by the gradient of the solution of the PDE.
- Optimal adversary is also characterized by the PDE.
- The hope is that we can solve the limiting PDE to have estimates on the growth of the regret and obtain explicit strategies for both parties.
- One can use the results of [Cover, 1966] and [Gravin et al., 2016] to solve this PDE for $K = 2$ and 3

$$u(x_1, x_2, x_3) = x^{(3)} + \frac{1}{2\sqrt{2}} e^{\sqrt{2}(x^{(2)} - x^{(3)})} + \frac{1}{6\sqrt{2}} e^{\sqrt{2}(2x^{(1)} - x^{(2)} - x^{(3)})}$$

where $x^{(1)} \leq x^{(2)} \leq x^{(3)}$; see [Drenska, 2017] and [Drenska and Kohn, 2019].
The advantage of the long time asymptotics

- For $\delta > 0$, the problem in hand is a game between the forecaster and the adversary.
- As $\delta \downarrow 0$ we only have a stochastic control problem for the adversary. The forecaster has to choose $\beta$ given by the gradient of the solution of the PDE.
- Optimal adversary is also characterized by the PDE.
- The hope is that we can solve the limiting PDE to have estimates on the growth of the regret and obtain explicit strategies for both parties.
- One can use the results of [Cover, 1966] and [Gravin et al., 2016] to solve this PDE for $K = 2$ and $3$

$$u(x_1, x_2, x_3) = x^{(3)} + \frac{1}{2\sqrt{2}}e^{\sqrt{2}(x^{(2)}-x^{(3)})} + \frac{1}{6\sqrt{2}}e^{\sqrt{2}(2x^{(1)}-x^{(2)}-x^{(3)})}$$

where $x^{(1)} \leq x^{(2)} \leq x^{(3)}$; see [Drenska, 2017] and [Drenska and Kohn, 2019].
The advantage of the long time asymptotics

- For $\delta > 0$, the problem in hand is a game between the forecaster and the adversary.
- As $\delta \downarrow 0$ we only have a stochastic control problem for the adversary. The forecaster has to choose $\beta$ given by the gradient of the solution of the PDE.
- Optimal adversary is also characterized by the PDE.
- The hope is that we can solve the limiting PDE to have estimates on the growth of the regret and obtain explicit strategies for both parties.
- One can use the results of [Cover, 1966] and [Gravin et al., 2016] to solve this PDE for $K = 2$ and 3

$$u(x_1, x_2, x_3) = x^{(3)} + \frac{1}{2\sqrt{2}}e^{\sqrt{2}(x^{(2)}-x^{(3)})} + \frac{1}{6\sqrt{2}}e^{\sqrt{2}(2x^{(1)}-x^{(2)}-x^{(3)})}$$

where $x^{(1)} \leq x^{(2)} \leq x^{(3)}$; see [Drenska, 2017] and [Drenska and Kohn, 2019].
The solution for 4 expert case

**Theorem (Bayraktar, E., Zhang (2020))**

\[
\begin{align*}
    u(x) &= x(4) - \frac{\sqrt{2}}{4} \sinh(\sqrt{2}(x(4) - x(3))) \\
    &\quad + \frac{\sqrt{2}}{2} \arctan \left( \frac{x(1) + x(2) - x(3) - x(4)}{e^{\sqrt{2}}} \right) \cosh \left( \frac{x(1) - x(2) + x(3) - x(4)}{\sqrt{2}} \right) \\
    &\quad - \frac{\sqrt{2}}{2} \arctanh \left( \frac{x(1) + x(2) - x(3) - x(4)}{e^{\sqrt{2}}} \right) \sinh \left( \frac{x(1) - x(2) + x(3) - x(4)}{\sqrt{2}} \right) \\
    &\quad + \frac{\sqrt{2}}{2} \arctan \left( \frac{-x(1) + x(2) + x(3) - x(4)}{e^{\sqrt{2}}} \right) \cosh \left( \frac{-x(1) - x(2) + x(3) + x(4)}{\sqrt{2}} \right) \\
    &\quad - \frac{\sqrt{2}}{2} \arctanh \left( \frac{-x(1) + x(2) + x(3) - x(4)}{e^{\sqrt{2}}} \right) \sinh \left( \frac{-x(1) - x(2) + x(3) + x(4)}{\sqrt{2}} \right)
\end{align*}
\]

- The comb strategy is asymptotically optimal for the adversary.
The solution for 4 expert case

Theorem (continued)

\[ u(x_1, x_2, x_3, x_4) = \frac{\pi}{4\sqrt{2}} + \frac{1}{4}(x_1 + x_2 + x_3 + x_4) + \]
\[ \frac{3\pi}{16\sqrt{2}}(x_1^2 + x_2^2 + x_3^2 + x_4^2) - \frac{2}{3}(x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4) \]
\[ + o(|x|^2). \]

Via Feynman-Kac, \( u \) can be written under the form

\[ u(x) = \sup_{\sigma \in \mathcal{A}} \mathbb{E} \left[ \int_0^\infty e^{-t} \Phi(X_t^{x,\sigma}) dt \right] \]

where \( \sigma \) is adapted to the filtration of a Brownian motion \( W \),

\[ dX_t^{x,\sigma} = \sigma_t dW_t \]

and \( \sigma_t = e^J \) for some \( J \subset \{1, \ldots, K\} \).
Parabolic problem

- We return to the deterministic stopping

\[ V(T, X_0) = \min_{\text{forecaster}} \max_{\text{adversary}} E^{\alpha,\beta}[\Phi(X_T)] \]

\[ = \max_{\text{adversary}} \min_{\text{forecaster}} E^{\alpha,\beta}[\Phi(X_T)]. \]

- We want to study the long time behavior \((T \to \infty)\) of \(V(T, 0)\). The correct rescaling is

\[ u^T(t, x) = \frac{1}{\sqrt{T}} V \left( T - tT, x\sqrt{T} \right) \text{ for } t \in [0, 1]. \]

- Then \(u^T\) converges to the unique viscosity solution with linear growth of

\[ \partial_t u(t, x) + \sup_{J \subset \{1, \ldots, K\}} \frac{1}{2} e_J^\top \partial_{xx} u(t, x) e_J = 0 \]

\[ u(1, x) = \Phi(x) \]
Parabolic problem cont’d

- For $K \leq 4$, $u(t, x)$ can be computed by inverse Laplace Transform of $u(x)$.
  - The growth of regret is $u(0, 0)\sqrt{T}$.
  - The optimal learning algorithm is $\partial_x u(t, x)$.
  - One can check the optimality of COMB strategy.

- The FvsG conjecture

\[
 u^T(0, 0) \sim \frac{2}{\sqrt{\pi}} u^{\delta=\frac{1}{4}}(0)
\]

- $K = 2$ solved by [Gravin et al., 2016], and $K = 3$ solved by [Abbasi Yadkori et al., 2017] using properties of random walks.

- In [Bayraktar, E., and Zhang, 2021], our method provides a general link between the finite time problem and geometric stopping and proves the conjecture for $K = 4$.

- We use tools from viscosity solutions and stochastic control (in continuous time!!).
Parabolic problem cont’d

- For $K \leq 4$, $u(t, x)$ can be computed by inverse Laplace Transform of $u(x)$.
  - The growth of regret is $u(0, 0) \sqrt{T}$.
  - The optimal learning algorithm is $\partial_x u(t, x)$.
  - One can check the optimality of COMB strategy.

- The FvsG conjecture

$$u^T(0, 0) \sim \frac{2}{\sqrt{\pi}} u^{\delta=\frac{1}{T}}(0)$$

- $K = 2$ solved by [Gravin et al., 2016], and $K = 3$ solved by [Abbasi Yadkori et al., 2017] using properties of random walks.

- $K = 2$ solved by [Gravin et al., 2016], and $K = 3$ solved by [Abbasi Yadkori et al., 2017] using properties of random walks.

- In [Bayraktar, E., and Zhang, 2021], our method provides a general link between the finite time problem and geometric stopping and proves the conjecture for $K = 4$.

- We use tools from viscosity solutions and stochastic control (in continuous time!!).
Parabolic problem cont’d

- For $K \leq 4$, $u(t, x)$ can be computed by inverse Laplace Transform of $u(x)$.
  - The growth of regret is $u(0, 0)\sqrt{T}$.
  - The optimal learning algorithm is $\partial_x u(t, x)$.
  - One can check the optimality of COMB strategy.

- The FvsG conjecture

\[
\begin{align*}
  u^T(0, 0) &\sim \frac{2}{\sqrt{\pi}} u^{\delta = \frac{1}{T}}(0) = \frac{u^{\delta = \frac{1}{T}}(0)}{2i\pi} \int_{x_0-i\infty}^{x_0+i\infty} e^{\lambda} \lambda^{-3/2} d\lambda.
\end{align*}
\]

- $K = 2$ solved by [Gravin et al., 2016], and $K = 3$ solved by [Abbasi Yadkori et al., 2017] using properties of random walks.

- In [Bayraktar, E., and Zhang, 2021], our method provides a general link between the finite time problem and geometric stopping and proves the conjecture for $K = 4$.

- We use tools from viscosity solutions and stochastic control (in continuous time!!).
Minimal regret wrt to $K$:

- For $K = 2$ it is $(2/\sqrt{\pi})(1/2)\sqrt{T/2}$,
- For $K = 3$ it is $(2/\sqrt{\pi})(2/3)\sqrt{T/2}$,
- For $K = 4$ it is $(2/\sqrt{\pi})(\pi/4)\sqrt{T/2}$.
- For $K = 5$ COMB strategy seems to be suboptimal [Chase, 2019].
For $K = 2$ it is $\left(\frac{2}{\sqrt{\pi}}\right)\left(\frac{1}{2}\right)\sqrt{T}/2$,

For $K = 3$ it is $\left(\frac{2}{\sqrt{\pi}}\right)\left(\frac{2}{3}\right)\sqrt{T}/2$,

For $K = 4$ it is $\left(\frac{2}{\sqrt{\pi}}\right)\left(\frac{\pi}{4}\right)\sqrt{T}/2$.

For $K = 5$ COMB strategy seems to be suboptimal [Chase, 2019].
Minimal regret wrt to $K$

- For $K = 2$ it is $(2/\sqrt{\pi})(1/2)\sqrt{T}/2$,
- For $K = 3$ it is $(2/\sqrt{\pi})(2/3)\sqrt{T}/2$,
- For $K = 4$ it is $(2/\sqrt{\pi})(\pi/4)\sqrt{T}/2$.
- For $K = 5$ COMB strategy seems to be suboptimal [Chase, 2019].
Limited adversary

- In [Bayraktar, E., and Zhang, 2021] we study the problem of prediction against limited adversary who corrupts the outcomes.
- Each expert has an exogenously given "success rate" and the adversary can corrupt only one expert. The problem bridges the adversarial and the stochastic settings.
- In the long-time regime, we obtain

\[
\partial_t u(t, x) + H(\partial_x u(t, x), \partial_{xx} u(t, x)) = 0 \\
u(1, x) = \Phi(x).
\]

- Fully nonlinear version of the PDE obtained in [Drenska and Calder, 2020].
- The PDE is similar to equations studied in [Chen et al., 1991] and [Soner et al., 2003], but neither geometric nor satisfies a comparison result.
- For the purposes of obtaining algorithms or obtaining the growth of regret, we only need to compare a smooth supersolution to a viscosity subsolution: partial comparison.
Outline

1 Introduction

2 Full Information
   - The Geometric stopping problem
   - Long time asymptotics
   - Our Main contribution
   - Finite vs. geometric regret conjecture

3 Partial Information Set-up
   - Derivatives in Wasserstein space
   - Limiting PDE
   - Verification approach
   - Viscosity approach
Set Up

- Both parties observe the same source of information which is the history of \((l_t, 1_{l_t \in J_t})\) and they know the distribution \(m_t\) of \(X_t\) conditional on their information.
- They choose their strategies \(\alpha_t \in A\) and \(\beta_t \in B\) and \(J_t, l_t\) are realized with distribution \(\alpha_t, \beta_t\).
- The (unobserved) gains of each experts selected by the adversary increases by 1 i.e.,
  
  \[
  G_t^i = G_{t-1}^i + 1 \text{ if } i \in J_t \\
  G_t^i = G_{t-1}^i \text{ if } i \not\in J_t.
  \]

- The observed gain of the forecaster is updated as
  
  \[
  G_t = G_{t-1} + 1 \text{ if } l_t \in J_t.
  \]

- The hidden state variable evolves as
  
  \[
  X_{t+1} - X_t := e_{J_t} - 1_{l_t \in J_t} e \in \mathbb{R}^K.
  \]
\[ Y_t := 1_{l_t \in J_t} - 1_{l_t \not\in J_t} \in \{ \pm i \} \]

indicates whether the forecaster makes a good decision or not. Both players can observe the law of adversary’s control \( a_{t-1} \) and the indicator \( y_{t-1} \). Their accumulated information is given by

\[ h_t := (m_0, a_0, y_0 \ldots, a_{t-1}, y_{t-1}) \in \mathcal{H}_t, \quad (h_0 := m_0 \in \mathcal{H}_0), \]

where \( \mathcal{H}_t := \mathcal{P}(\mathbb{R}^K) \times \mathcal{P}(\{0, 1\}^K) \times \{ \pm i \}^t \).

- The strategies of the forecaster and the adversary are measurable functions \( \beta_t : \mathcal{H}_t \to \mathcal{P}([K]) \) and \( \alpha_t : \mathcal{H}_t \to \mathcal{P}(\{0, 1\}^K) \) respectively. Strategies of the adversary \( \mathcal{A} \), of the forecaster \( \mathcal{B} \).
\[ Y_t := 1_{l_t \in J_t} I_{t - 1} l_t \not\in J_t I_{t - 1} \in \{\pm i\} \]

indicates whether the forecaster makes a good decision or not. Both players can observe the law of adversary’s control \( a_{t-1} \) and the indicator \( y_{t-1} \). Their accumulated information is given by

\[ h_t := (m_0, a_0, y_0 \ldots , a_{t-1}, y_{t-1}) \in \mathcal{H}_t, \quad (h_0 := m_0 \in \mathcal{H}_0), \]

where \( \mathcal{H}_t := \mathcal{P}(\mathbb{R}^K) \times (\mathcal{P}(\{0, 1\}^K) \times \{\pm i\})^t. \)

The strategies of the forecaster and the adversary are measurable functions \( \beta_t : \mathcal{H}_t \to \mathcal{P}([K]) \) and \( \alpha_t : \mathcal{H}_t \to \mathcal{P}(\{0, 1\}^K) \) respectively. Strategies of the adversary \( A \), of the forecaster \( B \).
Update of the conditional law of $X$

For any $\alpha \in \mathcal{A}$ and distribution $m$ on $\mathbb{R}^K$, denote

$$\hat{\alpha}(i) := \sum_{j,i \in j} \alpha(j), \quad \hat{\alpha}(-i) := \sum_{j,i \notin j} \alpha(j)$$

$$A_{i,\sqrt{T}}^{\alpha,m} = \sum_{j,i \in j} \frac{\alpha(j)}{\hat{\alpha}(i)} m_{\#_{\frac{e_j}{\sqrt{T}}}}^{\#_{\frac{e_j}{\sqrt{T}}}}, \quad A_{-i,\sqrt{T}}^{\alpha,m} = \sum_{j,i \notin j} \frac{\alpha(j)}{\hat{\alpha}(-i)} m_{\#_{\frac{e_j}{\sqrt{T}}}}^{\#_{\frac{e_j}{\sqrt{T}}}}$$

where $m_{\#_v}$ is defined by $\int f(x) dm_{\#_v}(x) = \int f(x + v) dm(x)$.

- By Bayes formula, the conditional distribution is updated as

$$m_{t+1} = 1_{\{l_t \in J_t\}} A_{l_t,1}^{\alpha,m_t} + 1_{\{l_t \notin J_t\}} A_{-l_t,1}^{\alpha,m_t}.$$
Upper value of the game

- We can iteratively define the upper value

\[
\overline{V}^T (s, m) = \inf_{\beta \in \mathcal{B}} \sup_{\alpha \in \mathcal{A}} \sum_i \beta(i) \alpha(i) \overline{V}^T \left( s + 1, A_{i,1}^{\alpha,m} \right) \\
+ \sum_i \beta(i) \alpha(-i) \overline{V}^T \left( s + 1, A_{-i,1}^{\alpha,m} \right)
\]

\[
\overline{V}^T (T, m) = \int \Phi(x) m(dx).
\]

- In this game, the forecaster announces his strategy, then the adversary announces a strategy, that might depend on the randomization of the forecaster.

- The update of the conditional distribution lacks convexity and the upper value might be different from the lower value.
Scaled upper value

- We define the scaled value function

$$u^T(s, m) := \frac{1}{\sqrt{T}} V \left( \lceil sT \rceil, m^*\sqrt{T} \right),$$

where $m^{*\lambda}$ is defined by $\int f(x)dm^{*\lambda}(x) = \int f(\lambda x)dm(x)$.

- It solves

$$u^T(s, m) = \inf_{\beta \in \mathcal{B}} \sup_{\alpha \in \mathcal{A}} \sum_i \beta(i)\alpha(i)u^T \left( s + \frac{1}{T}, A_{\alpha, \sqrt{T}}^{\alpha, m} \right) + \sum_i \beta(i)\alpha(-i)u^T \left( s + \frac{1}{T}, A_{-i, \sqrt{T}}^{\alpha, m} \right).$$

$$u^T(T, m) = \int \Phi(x)m(dx)$$

- If $u(t, m) = \lim_{T \to \infty} u^T(t, m)$, then at the leading order

$$V^T(0, \delta_0) = \sqrt{T}u(0, \delta_0) + o(\sqrt{T}).$$
Derivatives in Wasserstein space

- Let $u : \mathcal{M}_2(\mathbb{R}^K) \mapsto \mathbb{R}$ smooth enough. Using the notations of [Cardaliaguet et al., 2019, Chow and Gangbo, 2019, Carmona et al., 2018], we define $\frac{\delta u}{\delta m} : \mathcal{P}_2(\mathbb{R}^K) \times \mathbb{R}^K \mapsto \mathbb{R}$ by

$$
\lim_{h \to 0} \frac{u(m + h(m' - m)) - u(m)}{h} = \int \frac{\delta u}{\delta m}(m, x) d(m' - m)(x).
$$

and $D_m u(m, x) = D_x \frac{\delta u}{\delta m}(m, x)$ which is the sensitivity of the functional along push forward directions.

- We also define $\frac{\delta^2 u}{\delta m^2} : \mathcal{P}_2(\mathbb{R}^K) \times \mathbb{R}^K \times \mathbb{R}^K \mapsto \mathbb{R}$ as

$$
\frac{\delta^2 u}{\delta m^2}(m, x, y) = \frac{\delta}{\delta m} \left( \frac{\delta u}{\delta m}(m, x) \right)(y).
$$

- Then we can define $D^2_{mm} u(m, x, y) = D_x D_y \frac{\delta^2 u}{\delta m^2}(m, x, y)$ and

$$
\text{Hess} u(m) = D^2_{mm} u(m, [m], [m]) + D_x D_m u(m, [m]).
$$

where $f[m] = \int f(x) m(dx)$
Theorem

If \( u \) is \( C^1 \), then for all \( \alpha \in A \) and \( i \in \{1, \ldots, K\} \), we have

\[
\lim_{T \to \infty} \sqrt{T} \left( u(A_{\alpha, i}^T) - u(m) \right) = -\mathcal{V}_{\alpha, i}^\top D_m u(m, [m])
\]

and

\[
\lim_{T \to \infty} \sqrt{T} \left( u(A_{\alpha, -i}^T) - u(m) \right) = \mathcal{V}_{\alpha, -i}^\top D_m u(m, [m])
\]

where \( \mathcal{V}_{\alpha, i} = \sum_{j : i \in j} \frac{\alpha(j)}{\hat{\alpha}(i)} e_j \) and \( \mathcal{V}_{\alpha, -i} = \sum_{j : i \notin j} \frac{\alpha(j)}{\hat{\alpha}(-i)} e_j \).

The theorem states that at the leading order the impact of the (scaled)-Bayesian update on the value functional can be described using \( D_m u(m, [m]) \).
If $u$ is $C^2$, then for all $\alpha \in A$ and $i \in \{1, \ldots, K\}$, we have

$$
\lim_{T \to \infty} T \left( u(A_{\alpha, m}^\alpha, \sqrt{T}) - u(m) + \frac{1}{\sqrt{T}} V_{\alpha, i}^T D_m u (m, [m]) \right)
$$

$$
= \frac{1}{2} \sum_{j : i \in j} \frac{\alpha(j)}{\hat{\alpha}(i)} e_j^T D_x D_m u (m, [m]) e_j
$$

$$
+ \frac{1}{2} \sum_{k, j : i \in k, i \notin j} \frac{\alpha(j) \alpha(k)}{\hat{\alpha}(i) \hat{\alpha}(i)} e_j^T D_m^2 u (m, [m], [m]) e_k
$$

and

$$
\lim_{T \to \infty} T \left( u(A_{\alpha, m}^\alpha, -i, \sqrt{T}) - u(m) - \frac{1}{\sqrt{T}} V_{\alpha, -i}^T D_m u (m, [m]) \right)
$$

$$
= \frac{1}{2} \sum_{j : i \notin j} \frac{\alpha(j)}{\hat{\alpha}(-i)} e_j^T D_x D_m u (m, [m]) e_j
$$

$$
+ \frac{1}{2} \sum_{k, j : i \notin k, i \notin j} \frac{\alpha(j) \alpha(k)}{\hat{\alpha}(-i) \hat{\alpha}(-i)} e_j^T D_m^2 u (m, [m], [m]) e_k
$$
Theorem

The function

\[
\bar{u}(t, m) := \limsup_{(T, |t-t'| + W_2(m, m')) \to (\infty, 0)} u^T(t', m').
\]

is an USC viscosity subsolution of

\[
0 = \partial_t \bar{u}(t, m) + \frac{1}{2} \sup_{i, \alpha \in A} \text{Tr} \left( \text{Hess} \bar{u}(t, m) \Sigma (i, \alpha) + D_x D_m \bar{u}(t, m, [m]) \tilde{\Sigma} (i, \alpha) \right) \\
u(1, m) = \int \Phi(x)m(dx).
\]
Comparison to smooth supersolution

Theorem

Let $\phi$ be a smooth supersolution (with some growth condition) to

$$0 = \partial_t \phi(t, m) + \frac{1}{2} \sup_{i, \alpha \in A} \text{Tr} \left( \text{Hess} \phi(t, m) \Sigma (i, \alpha) + D_x D_m \phi(t, m, [m]) \tilde{\Sigma} (i, \alpha) \right)$$

$$\phi(1, m) = \int \Phi(x) m(dx)$$

then $\beta = D_m \phi(t, m, [m])$ is a mixed strategy for the forecaster leading to the bound

$$\limsup_{T \to \infty} \frac{V^T(0, 0)}{\sqrt{T}} \leq \phi(0, 0).$$

- Simplest candidate for supersolution $\phi(t, m) = \int f(t, x)m(dx)$ for some function $f(t, x)$. Then,

  $$D_m \phi(t, m, x) = D_x f(x), \quad D_x D_m \phi(t, m, x) = D^2_{xx} f(x), \quad D^2_{mm} \phi(t, m, x, y) = 0.$$  

- Thus, we need to find supersolutions to

  $$0 = \partial_t f(t, x) + \frac{1}{2} \sup_{i, \alpha \in A} \text{Tr} \left( D^2_{xx} f(t, x) \left( \tilde{\Sigma} (i, \alpha) + \Sigma (i, \alpha) \right) \right).$$
Comparison to smooth supersolution

**Theorem**

Let \( \phi \) be a smooth supersolution (with some growth condition) to

\[
0 = \partial_t \phi(t, m) + \frac{1}{2} \sup_{i, \alpha \in \mathcal{A}} \text{Tr} \left( \text{Hess}(\phi)(t, m) \Sigma(i, \alpha) + D_x D_m \phi(t, m, [m]) \tilde{\Sigma}(i, \alpha) \right)
\]

\[
\phi(1, m) = \int \Phi(x) m(dx)
\]

then \( \beta = D_m \phi(t, m, [m]) \) is a mixed strategy for the forecaster leading to the bound

\[
\limsup_{T \to \infty} \frac{V_T(0, 0)}{\sqrt{T}} \leq \phi(0, 0).
\]

- Simplest candidate for supersolution \( \phi(t, m) = \int f(t, x)m(dx) \) for some function \( f(t, x) \). Then,

\[
D_m \phi(t, m, x) = D_x f(x), \quad D_x D_m \phi(t, m, x) = D_{xx}^2 f(x), \quad D_{mm}^2 \phi(t, m, x, y) = 0.
\]

- Thus, we need to find supersolutions to

\[
0 = \partial_t f(t, x) + \frac{1}{2} \sup_{i, \alpha \in \mathcal{A}} \text{Tr} \left( D_{xx}^2 f(t, x) \left( \tilde{\Sigma}(i, \alpha) + \Sigma(i, \alpha) \right) \right).
\]
Simple bound

[B. Ekren, Zhang 2022]: There exists such an \( f \) and \( f(0, 0) = \sqrt{2 \ln(K)} \).
This leads to regret bounds
\[
\sqrt{2T \ln(K)}
\]
Similarly, using subsolutions we get a regret lower bound of the same order.
On going work: viscosity approach

- If we had absolutely continuous measures and non-degeneracy we could use [Gozzi and Świech, 2000].
- We want to prove comparison of viscosity solutions by using a smooth variational principle and the classical doubling of variable methodology, see [B., Ekren, and Zhang, 2022b].
- The equation of interest is the type of equations obtained in McKean-Vlasov optimal control problems with common noise. Also see [Soner and Yan, 2022, Cosso et al., 2021].

\[
\bar{u}(t, m) := \sup_{(I_s, \alpha_s)} \mathbb{E} \left[ \int \Phi(x) m^I_1,^A (dx) \right]
\]

where \( m^I_1,^A \) is a controlled version of flow of measure in [Chow and Gangbo, 2019] and \( (I_s, A_s) \in \{1, \ldots, k\} \times \mathcal{A} \) controls.
Thank you for your attention!
Near minimax optimal players for the finite-time 3-expert prediction problem.

Finite-time 4-expert prediction problem.

Prediction against a limited adversary.

A pde approach for regret bounds under partial monitoring.

A smooth variational principle on wasserstein space.
*Proceedings of the AMS*.

On the asymptotic optimality of the comb strategy for prediction with expert advice.

Bayraktar, E., Poor, H. V., and Zhang, X. (2021b).
Malicious experts versus the multiplicative weights algorithm in online prediction.


*The Master Equation and the Convergence Problem in Mean Field Games.*
Princeton University Press.


*Probabilistic Theory of Mean Field Games with Applications I-II.*
Springer.


How to use expert advice.


*Prediction, learning, and games.*


Experimental evidence for asymptotic non-optimality of comb adversary strategy.


Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations.
A partial laplacian as an infinitesimal generator on the wasserstein space.

Master bellman equation in the wasserstein space: Uniqueness of viscosity solutions.

Behavior of sequential predictors of binary sequences.
Technical report, STANFORD UNIV CALIF STANFORD ELECTRONICS LABS.

*A PDE Approach to a Prediction Problem Involving Randomized Strategies*.
ProQuest LLC, Ann Arbor, MI.
Thesis (Ph.D.)–New York University.

Online prediction with history-dependent experts: The general case.

Prediction with expert advice: A pde perspective.
*Journal of Nonlinear Science*. 

Towards optimal algorithms for prediction with expert advice.

Tight lower bounds for multiplicative weights algorithmic families.

A stochastic representation for mean curvature type geometric flows.

Viscosity solutions for mckean-vlasov control on a torus.