Mean-field-type learning is exactly what you need*

DARYL NOUPA YONGUENG[1], MAMADOU LAMINE DOUMBIA[2] and HAMIDOU TEMBINE[3]

***Abstract*— Distributed learning strategies in competitive multi agent systems with uncertainty and asymmetric resources is challenging, as gradient-based methods often fail to reach global optima, suffer from instability or poor convergence. We propose a mean-field-type learning (MFTL) framework that operates on agent distributions, enabling collective adaptation through game theoretic interactions. By unifying exploration and coordination, MFTL improves robustness and stability. The framework is evaluated in a competitive electric cab scenario of two companies with unequal fleet sizes and energy constraints, where simulations show smoother revenue dynamics and improved coordination over classical gradient-based methods.**

Keywords: Mean-Field Learning, Multi-Agent Systems, Game Theory, Distributed Optimization, Stochastic Dynamics.

I. INTRODUCTION

LEARNING optimal strategies in distributed and multi-agent systems remains a fundamental challenge in control and machine intelligence. Deterministic gradient-based methods often become trapped in local minima, while stochastic approaches with fixed noise levels can exhibit persistent instability, limiting convergence to globally optimal solutions. These issues are particularly acute in large-scale, interacting agent environments.

To address these limitations, we propose a mean-field-type learning (MFTL) framework that bridges deterministic and stochastic learning through a distributional formulation of the learning dynamics. Rather than updating only individual parameters, agents learn the evolution of their associated probability distributions, naturally incorporating interactions through the mean field. This collective adaptation promotes efficient exploration, mitigates variance-induced instability, and enables convergence beyond local extrema capabilities that standard methods such as ADAM or fixed-noise stochastic descent often lack. The effectiveness of the proposed approach is demonstrated through a competitive scenario involving two electric cab companies with unequal fleet sizes operating in the same urban environment.

Related works

The work in [1] analyzes Gibbs measures and shows concentration of probability mass on global minimizers. The asymptotic role of the spectral gap in simulated annealing convergence is studied in [2]. In [3], the authors propose swarm-based gradient methods in the Monge-Kantorovich space, with guarantees of convergence to global minimizers and extensions of simulated annealing via particle interactions. Convergence of Langevin simulated annealing with multiplicative noise in non-convex settings is established in [4]. Finally, [5] introduces "One swarm per Queen" algorithm for stochastic games based on inter-swarm communication, closely related to our mean-field-type learning framework.

Contribution

We introduce mean-field-type learning (MFTL), a distributed framework unifying deterministic and stochastic gradient dynamics at the distributional level. Inspired by swarm behavior, MFTL enables implicit exploration and coordination through mean-field interactions. We prove concentration of the stationary distribution on global minimizers, even for non-convex costs, with convergence rate guarantees. Unlike virtual-agent methods, MFTL operates directly on distributions, improving efficiency, stability, and reliability.

Table I displays the notations used in the manuscript.

Notation	Meaning
t	time
$\epsilon(t)$	vanishing function with time
θ	parameter
$c(\theta)$	cost function
c_i	cost at i
$\mu(t, d\theta)$	density of θ at time t also called mean-field
$L^1(\mathbb{R})$	the set of Lebesgue-integrable functions
$\mathcal{L}(\mathbb{R}^d, \mathbb{R}^k)$	set of linear operators from $\mathbb{R}^d$ to $\mathbb{R}^k$ represented by matrices in $\mathbb{R}^{d\times k}$
$\mathcal{P}(\mathcal{S})$	set of probability measures on $\mathcal{S}$
$\mathbb{P}_x$	the probability law of the random variable x
div_θ	$\sum_i \partial_{\theta_i}$
f_θ	$\partial_\theta f$
$f_{\theta\theta}$	$\partial_{\theta\theta} f$

TABLE I: Notations used in the manuscript

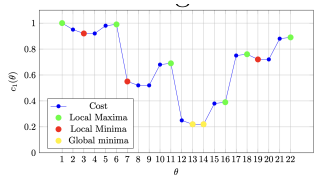
II. BOLTZMANN-GIBBS DISTRIBUTION

We start with the Boltzmann-Gibbs distribution that is behind the MFTL algorithm and show that Boltzmann-Gibbs distribution is concentrated only at the global minimizers. Let $\epsilon(t) > 0,\ \forall t \geq 0$ and consider the replicator equation

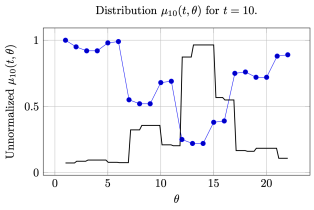
$$\dot{\mu}_i(t) = -\frac{1}{\epsilon(t)}\mu_i(t)\left(c_i - \sum_{j=1}^{I}\mu_j(t)c_j\right), \quad i \in \{1,\dots,I\}, t > 0 \tag{1}$$

*This work was supported by the U.S. Air Force Office of Scientific Research under Grant FA9550-17-1-0259

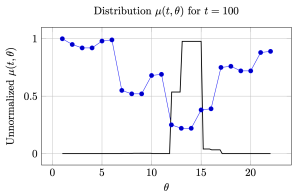
[1]DARYL NOUPA YONGUENG, MAMADOU LAMINE DOUMBIA, HAMIDOU TEMBINE, are with the Department of Electrical Engineering and Computer Engineering, University of Quebec at Trois-Rivieres, 3351, boulevard des Forges, G8Z 4M3, Canada daryl.noupa.yongueng@uqtr.ca



(a) Non-convex cost function



(b) Intermediary step of the algorithm



(c) Only global minima stay in the distribution

Fig. 1: 2D plots of $c_1(\theta) = (1,1)(2,0.95)(3,0.92)(4,0.92)(5,0.98)(6,0.99)(7,0.55)(8,0.52)(9,0.52)(10,0.68)(11,0.69)(12,0.25)(13,0.22)(14,0.22)(15,0.38)(16,0.39)(17,0.75)(18,0.76)(19,0.72)(20,0.72)(21,0.88)(22,0.89)$ and its MFTL finds the global minimizers at 12 and 13.

starting from the interior of the simplex:

$$\mu_i(0) = \mu_{i0} > 0, \quad \sum_{j=1}^{I} \mu_{j0} = 1 \tag{2}$$

It can be checked that

$$\mu_i(t) = \frac{\mu_{i0} e^{\left(-\int_0^t \frac{c_i}{\epsilon(t')}\, dt'\right)}}{\sum_{j=1}^{I} \mu_{j0} e^{\left(-\int_0^t \frac{c_j}{\epsilon(t'')}\, dt''\right)}} \tag{3}$$

solves the system. Moreover, as $\int_0^t \frac{1}{\epsilon(t')} dt'$ goes to infinity as t goes to infinity, the distribution $\mu(t)$ is concentrated at the set $\arg\min_{i\in\{1,\dots,I\}} c_i$ i.e., $support(\mu_i^*) \subseteq \arg\min_{i\in\{1,\dots,I\}} c_i$. Note that $\min_{i\in\{1,\dots,I\}} c_i = \min_{\{x\in\Delta_{I-1}\}} \sum_{j=1}^{I} x_j c_j$ where $\Delta_{I-1} = \{x \in \mathbb{R}^I \mid x_j \geq 0, \sum_{j=1}^{I} x_j = 1\}$ is the $(I-1)$-dimensional simplex of $\mathbb{R}^I$. This proves that the softargmin distribution (also called Boltzmann-Gibbs distribution) $\sim e^{\left(-(c_i - \min_k c_k)\int_0^t \frac{1}{\epsilon(t')}\, dt'\right)}$ converges to the set of global minimizers of $y \to \langle y, c\rangle$. Let $\mathcal{I}^* = \{i, c_i = \min_k c_k\} = \{i_1, i_2, \dots, i_k\}$. Then,

$$\lim_{t\to\infty} \mu_i(t) = \begin{cases} \frac{\mu_{i0}}{\sum_{j\in\mathcal{I}^*} \mu_{j0}} & \text{if } \ i \in \mathcal{I}^* \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

In learning, this means that the algorithm will tend to escape local minima, local maximum, or saddle points and converge towards the global optimum (see Figure 1 (a)-(b)-(c)). The gap between the Boltzmann-Gibbs distribution cost and the global minimum cost is

$$\sum_{i=1}^{I} \mu_i(t) c_i - \sum_{i=1}^{I} \mu_i^* c_* \to 0 \text{ as } \epsilon(t) \to 0$$

Example A typical example of application is in training a mean-field-type transformer with sigmoid self-attention. The obtained curve an Agent uses a mean-field-type transformer to make decision using the algorithm.

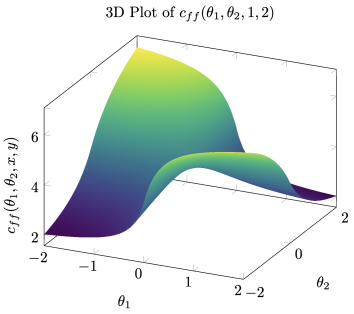


Fig. 2: 3D Plot of $c_{ff}(\theta_1,\theta_2,x,y)$

As illustrated in Figure 2 with the function $c_{ff}(\theta_1,\theta_2,x,y) = \left(\frac{\frac{\theta_1\theta_2 x}{1+|\theta_1 x|}}{1+|\frac{\theta_1\theta_2 x}{1+|\theta_1 x|}|} - y\right)^2$, the cost function $\theta \mapsto c(\theta,x,y)$ obtained from the transformer-based deep neural network, is generally non-convex. For the non-convex cost function $\theta \mapsto c(\theta,x,y)$ obtained from the transformer-based deep neural network, if the cost satisfies the above assumptions with finitely many global minima, the resulting Boltzmann-Gibbs measure $\mu(t,d\theta)$ will have a support concentrated only at its global minimizers. All the local minimizers and local/global maximizers and saddle points (if any) will not appear in the limiting support. We will use this property to build a mean-field-type learning algorithm for $\mu(t,d\theta)$ for both convex and non-convex functions. In Figure 1, subfigures 1b, 1c we plot the behavior of the unnormalized distribution $\mu(t,d\theta)$ when $t = 10$ and $t = 400$ with a properly designed Hurst index making a subdiffusion. We observe that the distribution concentrated very quickly to the set of global minima. Figure 1 illustrates a typical example where the classical gradient fail but the Boltzmann-Gibbs with high rationality is closer to the set of global minimizer of the non-convex function (displayed in Figure 1-(a)).

III. Building a Mean-Field-Type Learning Algorithm

The mean-field-type learning does not have to be stochastic depending on the tools in hand. One can work for example with Fokker-Planck-Kolmogorov equation in the simple cases to derive the evolution $\mu(t,d\theta)$ over the horizon $[0,T]$ and estimates asymptotics as T gets larger. However Boltzmann-Gibbs distribution can be extremely slow

if $\epsilon(t)$ is not properly designed. Additive and fractional Brownian motion yield stationary Gibbs measures $\mu_\epsilon(d\theta) \propto e^{-2c(\theta)/(\epsilon\sigma^2)}d\theta$, guaranteeing global concentration. Other processes lack this property. Fractional motion controls exploration via Hurst H: sub-diffusion ($H < 1/2$) allows slower cooling; super-diffusion ($H > 1/2$) accelerates.

A. Mean-field-type learning with additive Brownian motion

Start with a initial measure density $\mu(0, d\theta) = \mu_0(\theta)d\theta$ and the partial differential equation $\partial_t\mu(t,\theta) - \frac{div_\theta(\mu(t,\theta)c_\theta(\theta))}{\epsilon(t)} - \frac{\sigma^2}{2}trace(\mu_{\theta\theta}(t,\theta)) = 0$.

Using $trace(\mu_{\theta\theta}(t,\theta)) = div_\theta(\mu(t,\theta)(\log\mu(t,\theta))_\theta)$, the latter equation can also be rewritten as $\mu_t(t,\theta) - div_\theta[\mu(t,\theta)(\frac{c_\theta(\theta)}{\epsilon(t)} + \frac{\sigma^2}{2}(\log\mu(t,\theta)))_\theta] = 0$.

We now explain why this Fokker-Planck-Kolmogorov equation is related to the global minimization of c in the entire domain $\mathbb{R}^d$ even for a non-convex function c.

For $\epsilon(t) = \epsilon$ we connect the stationary distribution to the set of measure satisfies in the weak sense $div_\theta[\mu_*(d\theta)(\frac{c_\theta(\theta)}{\epsilon(t)} + \frac{\sigma^2}{2}(\log\mu_*(\theta)))_\theta] = 0$ which means that if μ is smooth the function $\frac{c(\theta)}{\epsilon} + \frac{\sigma^2}{2}\log\mu_*(\theta)$ is constant on every connected component of the set $\{\mu_* > 0\}$. There is a unique stationary distribution, it is positive everywhere and it is given by $\mu_\epsilon(d\theta) = C^{-1}e^{-\frac{2c(\theta)}{\epsilon\sigma^2}}d\theta,\ C = \int_\theta e^{-\frac{2c(\theta)}{\epsilon\sigma^2}}d\theta$. μ_ϵ is a global minimizer of $\mu \mapsto \int\mu(d\theta)[\frac{c(\theta)}{\epsilon} + \log\mu]$. One has

$$\epsilon\left(\int\mu_\epsilon(d\theta)[\frac{c(\theta)}{\epsilon} + \log\mu_\epsilon]\right) - \inf_\theta c(\theta) \to 0$$

as ϵ goes to infinity.

Proposition: The measure μ_ϵ is the global minimizer of $\mu \mapsto c_\epsilon(\mu) := \int\mu(d\theta)[\frac{c(\theta)}{\epsilon} + \frac{\sigma^2}{2}\log\mu(\theta)]$ even when c is non-convex.

It's used in [2] to prove convergence when $\epsilon(t)$ depends on time but with carefully designed $\epsilon(t)$.

Unfortunately the dependence in time is crucial in order to get a small error. We want $\epsilon(t)$ to be small as time increases so that the error gap is reduced. A typical examples are $\epsilon(t) = \frac{1}{(1+t)^{\frac{1}{3}}}$ or $\epsilon(t) = \frac{1}{\sqrt{\log(e+t)}},\ t > 0$. In [4] it is shown that the following mean-field-type learning with multiplicative Brownian motion:

$$\begin{array}{l}\mu_t(t,\theta)\\ -div_\theta[\mu(t,\theta)(\frac{\sigma\sigma' c_\theta(\theta)}{\epsilon(t)}\\ -(\sum_{j=1}^d\partial_j(\sigma(\theta)\sigma'(\theta))_{ij})_{i\in\{1,\dots,d\}})]\\ -\frac{1}{2}trace(\sigma(\theta)\sigma'(\theta)\mu(t,\theta))_{\theta\theta} = 0,\end{array} \tag{5}$$

converges to global minimizers of c even with c non-convex but coercive and have elliptic properties outside a certain compact set and c has finitely global minimizers $(\theta_i^*)_{i\in\mathcal{I}},\ |\mathcal{I}| < +\infty,$ with $c_{\theta\theta}(\theta_j^*)$ being positive matrix at each of the these global minimizers. The algorithms with multiplicative noise prove to be faster than the algorithm with constant additive noise. The measure $\mu(t, d\theta) \xrightarrow{\mathcal{L}} \mu^* = \frac{1}{\sum_{j\in\mathcal{I}}\frac{1}{\sqrt{det(c_{\theta\theta}(\theta_j^*))}}}\sum_{i\in\mathcal{I}}\frac{1}{\sqrt{det(c_{\theta\theta}(\theta_i^*))}}.\delta_{\theta_i^*}$ as $\epsilon(t) \to 0$.

B. Mean-field-type learning with fractional Brownian motion

1) Mean-field-type learning with penalty

We consider a generalized fractional mean-field-type learning algorithm governed by the following time-inhomogeneous partial differential equation:

$$\begin{array}{l}\mu_t(t,\theta) - div_\theta\left[\mu()\left(\frac{c_\theta(\theta)}{H\epsilon(t)t^{2H-1}\sigma^2} + \phi'_\theta(\mu)\right)\right] = 0,\\ \mu(0, d\theta) = \mu_0(\theta)d\theta,\end{array} \tag{6}$$

where $\mu(t,\theta)$ represents the probability density function at time t and position θ. The function $\phi : [0, +\infty) \to \mathbb{R}$ is a convex, twice-differentiable function with $\phi''(\theta) > 0$ and $\phi'(0) = -\infty$, ensuring strong regularization and preventing densities from becoming zero. The parameter $\epsilon(t) > 0$ is a time-varying regularization parameter, and $\mu(0, d\theta) = \mu_0(\theta)d\theta$ is the initial distribution, assumed to be absolutely continuous with respect to the Lebesgue measure and within the domain of the regularization functional. The parameters H and σ are constants related to the fractional Brownian motion.

This equation can be interpreted as a Wasserstein gradient flow, which provides a powerful framework for understanding the evolution of probability measures. To demonstrate this, we introduce the Gateaux derivative, which, when followed by a gradient, yields the variational derivative $\partial_{\theta\mu} = \nabla_\theta\frac{\delta}{\delta\mu}$. We apply this concept in the distributional sense.

Consider the regularized functional $c_{\epsilon(t)}(\mu)$:

$$c_{\epsilon(t)}(\mu) = \int\frac{c(\theta)}{H\epsilon(t)t^{2H-1}\sigma^2}\mu(d\theta) + \int\phi(\mu(\theta))d\theta. \tag{7}$$

The first term, $\int\frac{c(y)}{H\epsilon(t)t^{2H-1}\sigma^2}\mu(dy)$, represents a linear energy term (linear in the measure μ) associated with the cost function $c(y)$, scaled by the time-dependent regularization. Minimizing this term with respect to μ yields a constant c_{inf}. The Wasserstein gradient of this term is given by:

$$-\text{div}_\theta\left[\mu(t,\theta)\left(\frac{c_\theta(\theta)}{H\epsilon(t)t^{2H-1}\sigma^2}\right)\right], \tag{8}$$

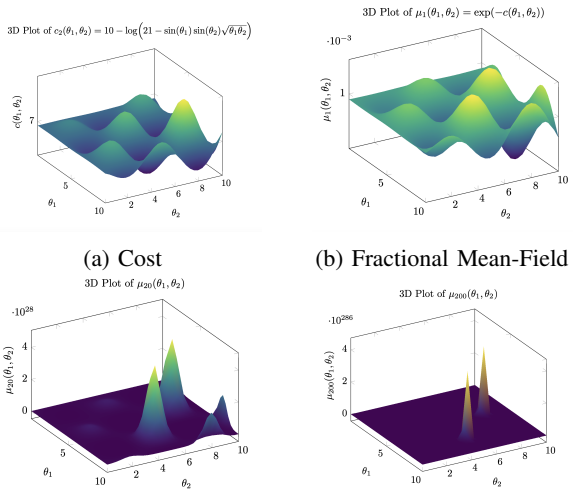
which is obtained by identifying the derivative in the sense of distributions.

The second term, $\int\phi(\mu(\theta))d\theta$, represents a regularization term that penalizes deviations from uniform distributions. Its Wasserstein gradient is:

$$-\text{div}_\theta\left[\mu(\theta)\phi'_\theta(\mu(\theta))\right], \tag{9}$$

While traditional gradient descent in $\mathbb{R}^d$ may fail to reach global minimizers, especially for non-convex functions, Wasserstein gradient flows in the space of probability measures can often achieve them. This is because the Wasserstein metric accounts for the geometry of probability distributions, allowing for more robust convergence properties. Therefore, fractional mean-field-type learning, when formulated as a Wasserstein gradient flow, offers a promising approach to reach global optima in both convex and non-convex settings. The regularization term ϕ and the time varying regularization

parameter $\epsilon(t)$ are crucial to guarantee the convergence and the well-definition of the solution.



(a) Cost (b) Fractional Mean-Field

(c) Fractional Mean-Field (d) Fractional Mean-Field

Fig. 3: 3D Plot of $c_2(\theta_1,\theta_2) = 10 - \log\Big(21 - \sin(\theta_1)\sin(\theta_2)\sqrt{\theta_1\theta_2}\Big)$

C. Building a multi-agent mean-field-type learning algorithm

We consider I machine intelligence agents, each agent i equipped with a transformer-based neural network parameterized by weights and biases θ_i. The cost function for agent i, denoted $c_i(\theta_1,\ldots,\theta_I,x,y)$, maps the collective network parameters to a scalar cost, reflecting the interdependence of agent outputs. The objective for agent i is to determine θ_i that minimizes c_i, specifically:

$$\begin{aligned}&\theta_i \in \arg\min_{\theta_i'\in\mathbb{R}^d} c_i(\theta_1,\ldots,\theta_{i-1},\theta_i',\theta_{i+1},\ldots,\theta_I,x,y),\\ &\quad \forall i\in\mathcal{I}=\{1,\ldots,I\},\end{aligned} \tag{10}$$

The increasing prevalence of collaborative or simultaneous task execution by multiple agents necessitates the exploration of inter-agent interactions. Given the non-convex nature of cost functions derived from transformer-based neural networks with respect to θ_i, standard gradient descent methods are rendered sub-optimal.

D. Individual mean-field-type learning

We consider a generalized fractional mean-field-type learning algorithm governed by the following time-inhomogeneous partial differential equation:

$$\begin{aligned}&\mu_i(0,d\theta_i) = \mu_{i0}(\theta_i)d\theta_i,\\ &\mu_{it}(t,\theta_i) - div_{\theta_i}\left[\mu_i(t,\theta_i)\left(\tfrac{c_{i,\theta_i}(\theta_1,\ldots,\theta_I,x,y)}{H\epsilon(t)t^{2H_i-1}\sigma_i^2}+\right.\right.\\ &\phi'_{i,\theta_i}(\mu_i(t,\theta_i)) = 0,\\ &i\in\mathcal{I}=\{1,\ldots,I\},\end{aligned} \tag{11}$$

In the case where there exists a single function $V(\theta_1,\ldots,\theta_I,x,y)$ such that $\arg\min_{\theta_i} c_i = \arg\min_{\theta_i} V$ and $c_{i,\theta_i}(\theta_1,\ldots,\theta_I,x,y) = V_{\theta_i}(\theta_1,\ldots,\theta_I,x,y)$ for all i one can relate to the best response of the potential game with cost $V(\theta_1,\ldots,\theta_I,x,y)$. Global minimizers of V are also equilibria of the game and the mean-field-type learning algorithm can be used to reach them. Note that the difference with the literature is it includes a class of non-convex potential functions.

E. Collective mean-field-type learning

We consider a collaborative mean-field-type learning algorithm with $c=\sum_{j=1}^I c_j$ governed by the following time-inhomogeneous partial differential equation:

$$\begin{aligned}&\mu(0,d\theta) = \mu_0(\theta)d\theta,\\ &\mu_t(t,\theta) - div_\theta\left[\mu(t,\theta)\left(\tfrac{c_\theta(\theta_1,\ldots,\theta_I,x,y)}{H\epsilon(t)t^{2H-1}\sigma^2}+\right.\right.\\ &\phi'_\theta(\mu(t,\theta)) = 0,\\ &i\in\mathcal{I}=\{1,\ldots,I\},\end{aligned} \tag{12}$$

IV. Numerical Example

As a numerical illustration of the proposed MFTL algorithm, we consider a distributed power network (DIPONET) with prosumers (producer-consumers) who charge their electric vehicles (EVs) while two competing electric cab companies operate in the same town. The objective of the cab companies is to optimize revenue while preserving battery life through effective management of charging and deployment decisions.

A. Battery State-of-Charge Dynamics

The battery state-of-charge (SOC) quantifies the energy stored relative to capacity. Its dynamics depend on charging/discharging currents, efficiencies, and self-discharge. For a battery of type j with capacity $C^{(j)}$, the normalized SOC $\text{SOC}^{(j)}(t)\in[0,1]$ evolves as

$$\begin{aligned}\text{SOC}^{(j)}(t) = \text{SOC}^{(j)}(t_0)+\\ \int_{t_0}^t \frac{1}{C^{(j)}}\eta_{\text{ch}}(T(t'))I_{\text{ch}}(t')\mathbb{I}_{\text{ch}}(t')dt'\\ -\int_{t_0}^t \frac{1}{C^{(j)}}\frac{1}{\eta_{\text{dr}}(T(t'))}I_{\text{dr}}(t')\mathbb{I}_{\text{dr}}(t')dt' -\\ \int_{t_0}^t \lambda(T(t'))\text{SOC}^{(j)}(t')dt',\end{aligned} \tag{13}$$

where $\eta_{\text{ch}}(T)$ and $\eta_{\text{dr}}(T)$ are temperature-dependent charge/discharge efficiencies, and $\lambda(T)$ is the self-discharge rate. Battery degradation is modeled through calendar and cycle aging. For capacity $C^{(j)}(t)$ of type j, the dynamics are

$$\frac{dC^{(j)}(t)}{dt} = -\left[k_{\text{cal}}^{(j)}(T(t)) + k_{\text{cyc}}^{(j)}(T(t))r(t)\right], \tag{14}$$

with $k_{\text{cal}}^{(j)}$ given by an Arrhenius law and $k_{\text{cyc}}^{(j)}$ dependent on depth-of-discharge and operating current. The remaining useful life (RUL) is defined as the minimum time τ until the state of health reaches an end-of-life threshold αC_0.

B. Seasonal Operation Scenarios

We examine two representative regions: 1) Four-season climate with winter ($-40°$C, $-10°$C), spring/fall (mild), and summer ($25°$C, $32°$C), each with distinct EV usage frequencies. 2) Two-season tropical climate with rainy (mild) and hot (high-degradation) periods. Cycle frequencies $r(t)$ and temperatures $T(t)$ are piecewise constant by season, and capacity degradation is integrated annually to estimate RUL.

C. Data-Driven and Application to Competing Cab Companies

To improve remaining useful life (RUL) estimation, we integrate a knowledge-based transformer trained on the NASA battery dataset [6], embedding electrochemical degradation laws into the learning architecture for improved accuracy and interpretability. Within the DIPONET framework, competing cab companies apply mean-field-type learning (MFTL) to jointly optimize fleet deployment and charging decisions, balancing short term revenue with long-term battery health under seasonal variations and heterogeneous battery chemistries.

1) Battery Specifications and Empirical SOH Data

To validate the proposed model, we consider 4 commercial lithium-based cells with publicly available performance data. These include: (i) a Li-ion polymer rechargeable cell, and (iv) a high-power APR18650M1A cell. The specification sheets report charge/discharge curves, cycle life, and temperature-dependent discharge characteristics [7]–[10]. Representative results are shown in Figs 4a, 5b. These datasets highlight the impact of depth-of-discharge, C-rate, and ambient temperature on state-of-health (SOH) degradation.

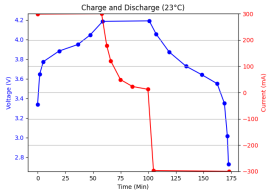
2) SOH Modeling Framework

Our degradation model of battery behavior is modeled using empirical manufacturer data capturing C-rate ($0.5C$, $1C$) and temperature ($-50°$C to $25°$C). Figures 6a illustrate the model predicted degradation trajectories.

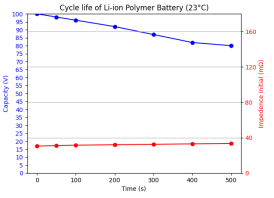
3) SOH of 2 batteries : data from some manufacturers

Li-ion Polymer rechargeable Battery specification sheet

This specification curves describe the basic performance, technical parameters, testing methods [7]. Figure 4a displays charge and discharge curve of Li-ion Polymer rechargeable Battery. Figure 4b expresses the cycle life of the Li-ion polymer battery at $(23 \pm 2^0 C)$.



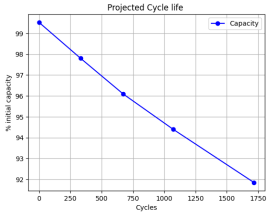
(a) Charge and discharge $(23 \pm 2°C)$ of Li-ion Polymer rechargeable Battery



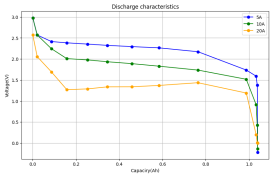
(b) Cycle Life $(23 \pm 2°C)$ of Li-ion Polymer rechargeable Battery

High Power Lithium Ion APR18650M1A

A123Systems' lithium ion rechargeable APR18650M1A cell is capable of very high power, long cycle and storage life, and has superior abuse tolerance due to the use of patented $Nanophosphate^{TM}$ technology [8]. Figure 5a displays the discharge characteristics, room temperature, the currents are $5A$, $10A$, $20A$. Figure 5b represents the projected cycle life 100% DOD, 1C/1C, Room temperature.



(a) Discharge characteristic by room temperature for A123Systems' lithium ion rechargeable APR18650M1A cell.

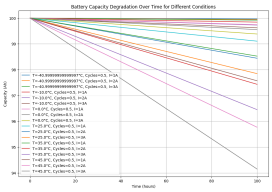


(b) Projected cycle life at 100% DOD, 1C/1C, room temperature for A123Systems' APR18650M1A cell.

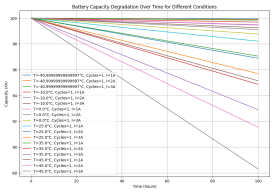
Fig. 5: Performance characteristics of A123Systems' lithium-ion rechargeable APR18650M1A cell.

4) SOH of batteries : model

Below, the plot of degradation from our model depends on the cycle rate, the temperature and the current intensity. We consider -50^0C, -10^0C, 0^0C, 10^0C and 25^0C for the temperature. The number of cycle 0.5C, 1C and the intensity of charge 1A, 2A, 3A. Figures 6a, 6b below are illustrations of the curve from our model for 0.5 cycle, 1 cycle. We observe that the model captures the empirical data provided by the manufacturers.



(a) Degradation for 0.5C generated by our model.



(b) Degradation for 1C generated by our model.

Fig. 6: Degradation curves generated by our model for different C-rates.

5) Impact on RUL and Cab Company Application

A trivial optimization would minimize the cycle rate $r(t)$ to preserve RUL, yielding $r^*(t) = r_{\min} \geq 0$. However, such a strategy fails to meet operational demand. In the cab company competition scenario, a trade-off emerges: maximizing fleet service revenue requires higher utilization, while preserving RUL favors conservative operation. The proposed MFTL framework is then employed to balance these conflicting objectives by coordinating deployment and charging strategies across fleets.

Electric Cab Company Profit

Consider $I \geq 2$ cab companies operating in a dense city. Each company $i \in \mathcal{I} = \{1, \ldots, I\}$ operates n_i electric cabs, each equipped with two batteries managed by an AI system monitoring SOC, SOH, and RUL. Due to intensive usage, battery degradation is continuously tracked for predictive maintenance. Each company aims to maximize its total profit. For company i :

$$\Pi_{i,\text{total}} = \overbrace{\left(\sum_{t=1}^{T_i=7\times365} \sum_{j=1}^{n_{i,t}} \sum_{k=1}^{D_{ijd,t}} F_{ijkd,t} \mathbb{I}_{\{E_{ijkd,t} \geq E_{min}\}} \cdot (1 - tax) \right)}^{\text{7-year Revenue}} - \underbrace{\left[\left(\sum_{t=1}^{T_i=7\times365} \sum_{j=1}^{n_{i,t}} \sum_{k=1}^{D_{ijkd,t}} p_{e,t} \cdot E_{ijkd,t} + \frac{c_{r,ijk,t}}{RUL_{ijk,t}} \right) + \sum_{d=1}^{2} \sum_{j=1}^{n_{i,t}} S_{ijkd,t} \right]}_{\text{Operating Costs}} - \underbrace{(n_i \cdot (C_c + 2C_b)(1 + tax))}_{\text{Initial Costs}} \tag{15}$$

A simplified annual profit of the first year is:

$$\Pi_{i,\text{total}} = \underbrace{\left(D_i \cdot F \mathbb{I}_{\{E \geq E_{min}\}} \cdot (1 - t) \cdot 365 \right)}_{\text{Annual Revenue}} - \underbrace{\left[\left(D_i \cdot p_e \cdot E + \frac{D_i \cdot c_r}{RUL_i} \right) \cdot 365 + 2n_i \cdot S \right]}_{\text{Operating Costs}} - \underbrace{(n_i \cdot (C_c + 2C_b))}_{\text{Initial Costs}} \tag{16}$$

n_i: Number of electric cabs ($n_1 = 10{,}000$, $n_2 = 20{,}000$), etc. D: Total daily market demand (rides/day), $D_i = \frac{n_i^\alpha}{n_1^\alpha + n_2^\alpha + \ldots + n_I^\alpha} \cdot D$: Daily rides served by company i, α is a positive parameter, F: Fare per ride (\$), t: Tax rate (decimal), p_e: Electricity price (\$/kWh), E: Energy consumption per ride (kWh), c_r: Battery replacement cost (\$), RUL_i: Remaining useful life of batteries (cycles), S: Annual salary per driver (\$), C_c: Initial cost per cab (\$), C_b: Cost per battery (\$). $\mathbb{I}_{\{E_{ijkd,t} \geq E_{min}\}}$ indicates the condition that there is a min energy in the electrical vehicle to serve the customer for the ongoing trip. High upfront costs drive early losses, with battery degradation (RUL_i) and driver salaries dominating expenses. Profitability improves as capital costs are amortized. Company profits are interdependent through customer demand D, electricity price p_e, and driver competition, leading to a Nash equilibrium problem.

$$\begin{aligned} &\max_{n_i} \Pi_{i,\text{total}}(n) \\ &i \in \mathcal{I} = \{1, 2, \ldots, I\}. \end{aligned} \tag{17}$$

Figure 7 represents the profit of companies 1 and 2 for the following parameters: $F = 20\$$, $t = 0.2$, $P_e = 0.2$ \$/kWh, $E = 0.5$, $c_r = 0.00001$, $C_c = 73000$, $S = 50000, C_b = 20000, D = 164383, RUL = 8000, \alpha = 0.025$. Figure 8 finds the corresponding equilibria using MFTL. This application uses fractional Brownian motion ($H < 1/2$) to enhance exploration in non-convex, multi-modal cost landscapes (Figures 3).

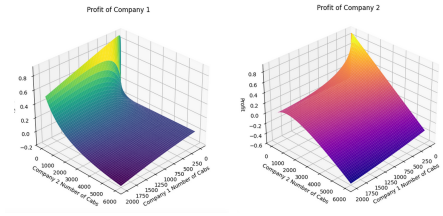


Fig. 7: Profits of Companies.

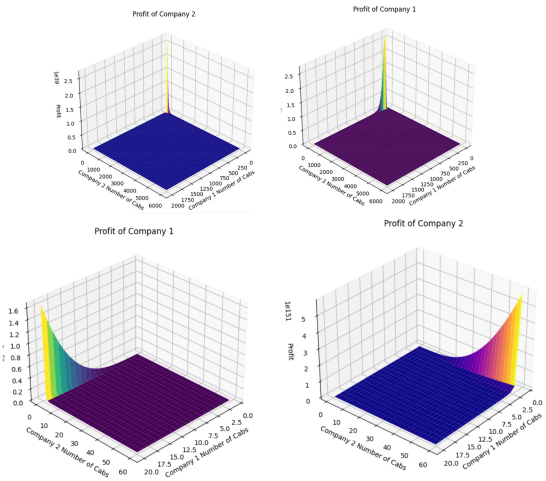


Fig. 8: Mean-field-type Nash equilibria.

V. Conclusion

This paper proposes an adaptive method of multi-agent machine intelligence: a mean-field-type learning for finding global minima and mean-field-type equilibria in non-convex cost functions. We validated the method on battery RUL in competing cab fleets. Future work will include risk-sensitive expectile agents.

References

[1] Krishna B. Athreya and Chii-Ruey Hwang. Gibbs measures asymptotics. Sankhya A, 72(1):191-207, February 2010.

[2] Holley R., Kusuoka, S., Stroock D.: Asymptotics of the spectral gap with applications to the theory of simulated annealing. J. Funct. Anal. 83(2), 333-347 (1989).

[3] Jérôme Bolte, Laurent Miclo, Stéphane Villeneuve, Swarm gradient dynamics for global optimization: the mean-field limit case, Mathematical Programming (2024), 205:661-701.

[4] Pierre Bras and Gilles Pagès, Convergence of Langevin-Simulated Annealing algorithms with multiplicative noise, Math. Comput.,2021, vol. 93, pp. 1761-1803,

[5] A. Tcheukam and H. Tembine, "One Swarm per Queen: A Particle Swarm Learning for Stochastic Games," 2016 IEEE 10th International Conference on Self-Adaptive and Self-Organizing Systems (SASO), Augsburg, Germany, 2016, pp. 144-145.

[6] NASA Ames Prognostics Center of Excellence (PCoE), 2022, available at: https://data.nasa.gov/dataset/Li-ion-Battery-Aging-Datasets/uj5r-zjdb/about_ data

[7] HONCELL, *Li-ion Polymer Rechargeable battery specification sheet*,Honcell Energy, 2022-05-07. www.honcell.com

[8] High Power Lithium Ion APR18650m1A, https://www.ecolithiumbattery.com/

[9] LTO18650 LIR18650 2600mAH, Lithium-ion battery, http://eemb.com

[10] High Power Lithium Ion APR18650m1A, www.a123systems.com