REPORT DOCUMENTATION PAGE

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
15-11-2012	Final Technical	01/09/2009 to 31/08/2012

4. TITLE AND SUBTITLE
Title of Grant:
Nonlinear Markov control processes and games

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA9550-09-1-0664

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
PI and Co-PIs
Kolokoltsov, Vassili N.

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The University of Warwick

8. PERFORMING ORGANIZATION REPORT NUMBER
R.STAA.3024

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR
875 N. Randolph St
Arlington, VA 22203

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Choose one:
A = Approved for public release; distribution is unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The project was devoted to the analysis of a new class of stochastic games, nonlinear Markov games, as they arise as a (competitive) controlled version of nonlinear Markov processes, which can be characterized by the property that the future depends on the past via present position and its distribution. This class of games can model a variety of situations for economics and epidemics, statistical physics and pursuit - evasion processes. Another key motivation arises from the steady increase in complexity of the modern technological development requires an appropriate (or better optimal) management of complex stochastic systems consisting of large number of interacting components (agents, mechanisms, vehicles, subsidiaries, species, police units, etc) , which may have competitive or common interests. Under rather general assumptions, the limiting problem as the number of components tends to infinity can be described by a nonlinear evolution on measures, and its controlled version is given precisely by a nonlinear Markov control process or (in case of competitive interests) a nonlinear Markov game that we are investigating.

15. SUBJECT TERMS
nonlinear Markov game, nonlinear Markov control, models of deception, complex stochastic systems, interacting agents

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
unclassified	unclassified	unclassified	unclassified (U)	22	PI name: Professor Kolokoltsov
					19b. TELEPHONE NUMBER 442476574806

NONLINEAR MARKOV CONTROL PROCESSES AND GAMES: FINAL REPORT

Vassili N. Kolokoltsov

Abstract

The project was devoted to the analysis of a new class of stochastic games that I called nonlinear Markov games, as they arise as a (competitive) controlled version of nonlinear Markov processes, which can be roughly characterized by the property that the future depends on the past not only via the present position (as in usual Markov processes), but also via its distribution. This class of games can model a variety of situations for economics and epidemics, statistical physics and pursuit - evasion processes.

Nonlinear Markov games can be considered as a systematic tool for modeling deception. In particular, in a game of pursuit - evasion, an evading object can create false objectives or hide in order to deceive the pursuit. Thus, observing this object leads not to its precise location, but to its distribution only, implying that it is necessary to build competitive control on the basis of the distribution of the present state. Moreover, by observing the action of the evading objects, one can make conclusions about its certain dynamic characteristics making the (predicted) transition probabilities depending on the observed distribution via these characteristics. This is precisely the type of situations modeled by nonlinear Markov games.

Another key motivation arises from the steady increase in complexity of the modern technological development requires an appropriate (or better optimal) management of complex stochastic systems consisting of large number of interacting components (agents, mechanisms, vehicles, subsidiaries, species, police units, etc) , which may have competitive or common interests. Carrying out a traditional Markov decision analysis for a large state space is often unfeasible. However, under rather general assumptions, the limiting problem as the number of components tends to infinity can be described by a well manageable nonlinear deterministic evolution on measures, and its controlled version is given precisely by a nonlinear Markov control process or (in case of competitive interests) a nonlinear Markov game that we are investigating.

The results of the project concern the fundamental mathematical questions of the theory of nonlinear Markov control processes and games like well posedness and controllability, as well as more applied issues such as convergence of approximating schemes. The latter are linked with interacting particle approximations, as introduced above.

1 Objectives for each grant year

The overall aim of the project was to address both the fundamental questions of the theory of nonlinear Markov control processes and games like well posedness and controllability, and the more applied issues such as approximation and numeric schemes.

Simple illustrative examples to have in mind were: (1) Pursuit - evasion: the evader produces false targets, so that the control of the pursuer should be based not on the observed position of the evader, but only on the observed distribution of this position; (2) Finances: observing performance of the competitor-company allows one to make a conclusion on the distribution of certain hidden internal parameters of this company, and hence to make decisions based on this distribution. (3) Similarly, the traces of the actions of terrorists or other organize crime groups can be used to assess the probability distribution of their actual states (physical locations, amount of equipment available, etc.), which again leads to the problem of control on the basis of the knowledge of the probability law on the state space thus relating nonlinear Markov control processes to the methods of crime (say, terrorist attacks) prevention.

Another crucial point for modern modeling in finance or inspection - crime prevention measures is in making decision on the basis of certain risk characteristics like variance or VaR (Value at Risk), which represent functions of the whole distribution, and not only on the position of a process at a given time.

Let us state the objectives for each year.

Tasks for Year 1.

Before plunging seriously into the control setting, the analysis of nonlinear Marov processes themselves was to be developed starting with the simplest classes such as nonlinear Levy processes and nonlinear Markov chains. The analysis had to include basic constructions, well-posedness issues, qualitative behavior and approximating schemes.

To pave the way for possibly wider applications, the links with concrete problems of natural science should be explicitly established including the models of non-equilibrium statistical mechanics, the replicator dynamics of multi agent evolutionary games (evolutionary biology), relevant models of financial dynamics and disease spreading.

Tasks for Year 2.

The main core of the research proposed is in the developing of the theory of nonlinear Markov processes and their controlled versions including competitive control. Initiated in year 1 mostly on the level of discrete models, this task had to be fully completed in Year 2.

Tasks for year 3.

The main problem is in linking the theoretical construction of nonlinear Markov processes with controlled system of interacting particles bringing discrete approximation with algorithmic methods of numeric calculations and more concrete applied models, like decision making or controlling large robot swamps or large armies. Thus we have to establish a rigorous link with two-sided applicability. Firstly, in order to be able to apply the theoretic results to concrete models of practical interest, the numeric schemes for the solutions are to be developed together with appropriate estimates for error terms. The most natural approximation and related algorithms are based on the approximations by systems of a large number of interacting particles. On the other hand, solving limiting nonlinear control Markov process can lead to a useful qualitative and quantitative asymptotics to the system of interacting particles.

As motivation for further research we indicated possible extensions to state spaces with nontrivial geometry, to the controlled nonlinear quantum dynamic semigroups and related nonlinear quantum Markov processes, as well as to the full infinite-dimensional measure valued control Markov processes and games.

2 Findings for each objective

2.1 Findings for year 1

The large part of the work for Year 1 was devoted to discrete models.

We have summarized the nonlinear analogues of the basic theory of usual Markov chains, where measures (on a finite state space) are described by a finite-dimensional simplex. A discrete space nonlinear Markov semigroup is a one-parameter semigroup of (possibly nonlinear) transformations of the unit simplex in n-dimensional Euclidean space (which represents the set of probability laws in a finite set of n points. In stochastic representation these transformations are given by stochastic matrices (as for usual Markov chains) depending on a position (nonlinearity!), whose elements specify nonlinear transition probabilities. Our first result yields the nonlinear analog of the basic convergence to a stationary regime from the theory of Markov chains, basic conditions being certain mixing property of nonlinear transition probabilities. In case of the semigroup parametrized by continuous time one defines its generator as the derivative of the semigroup at time zero. Stochastic representation for the generator means its representation by a Q-matrix (or infinitesimally stochastic matrix) again depending on a position. Examples are numerous: replicator dynamics, Lottka-Volterra model, basic epidemics, see [1].

For the corresponding control process we obtain nonlinear analogs of the basic long time behavior result, showing the existence of the limiting average income per unit of time and of the stationary strategies (turnpikes), see [1]. Related results were later developed in [9] on a somewhat more systematic and general grounds that are not yet fully exploited in the nonlinear case.

Let us point out the (not so obvious) place of the usual stochastic control theory in this nonlinear setting. Namely, even assuming that the transition probabilities do not depend on the distribution, does not reduce the problem to the usual stochastic control setting, but to a game with incomplete information, where the states are probability laws. That is, when choosing a move the players do not know the position precisely, but only its distribution.

The analysis of nonlinear Markov processes was systematically developed from two complimentary points of view: (i) analytic, based on functional analytic technique of semigroups and operators, where the main object was the nonlinear kinetic equation in the weak form of the type

$$\frac{d}{dt}(f,\mu_t) = (A_{\mu_t} f, \mu_t) \tag{1}$$

for the flow of Borel measures μ_t in $\mathbf{R}^d$, with a family of pseudo-differential generators of Markov processes of the type

$$L_\mu f(x) = \frac{1}{2}(G(x,\mu)\nabla,\nabla)f(x) + (b(x,\mu),\nabla f(x))$$
$$+\int (f(x+y)-f(x)-(\nabla f(x),y))\nu(x,\mu;dy),$$

see [1],[4]; and (ii) probabilistic, based on the related to (1) differential equations driven by nonlinear Lévy noise:

$$dX(t) = dY_t(X(t),\mathcal{L}(X(t))) \tag{2}$$

in $\mathbf{R}^d$ ($\mathcal{L}(X)$ denotes the probability law of the random variable X), where $Y_t(z,\eta)$ is a family of Lévy processes specified by the Lévy-Khinchine generators

$$L[z,\eta]f(x) = \frac{1}{2}(G(z,\eta)\nabla,\nabla)f(x) + (b(z,\eta),\nabla f(x))$$

$$+\int (f(x+y) - f(x) - (\nabla f(x), y))\nu(z,\eta;dy),$$

depending on a point z and a probability measure η in $\mathbf{R}^d$ as on parameters, see [1], [3], [7] (and some details and complements in [2]). The construction is given explicitly via the nonlinear analog of the Ito-Euler approximation scheme. This scheme also supplies the numeric algorithm for the practical calculations of the solutions.

The links with non-equilibrium statistical mechanics, the replicator dynamics of multi agent evolutionary games and epidemiology were established in [1] (evolutionary biology), financial models were given received even more attention highlighted in [5] and [10]. The examples again are numerous, as these evolutions exhaust all positivity preserving evolutions on measures subject to certain mild regularity assumptions. In particular, they include the Vlasov, Boltzmann, Smoluchovski, Landau-Fokker-Planck equations, as well as McKean diffusions and many other models.

Extending the link with usual Markov chains (described above) to general Markov processes with continuous state space, we can stress that, for a nonlinear Markov process, the future depends on the past not only via its present position, but also via its present distribution. A nonlinear Markov semigroup can be considered as a nonlinear deterministic dynamic system, though on a weird state space of measures. To give it a probabilistic interpretation one should specify a stochastic representation for this semigroup in terms of nonlinear transition probabilities satisfying the nonlinear analog of the Chapman-Kolmogorov equation, see in more details in Section 'Findings for year 2' below.

2.2 Findings for year 2

In Year 1 the theory of nonlinear Markov processes was developed for discrete state space and initiated for the general case. In Year 2 we completed this development.

Let us describe in more detail the central object of our study: a nonlinear Markov process. Loosely speaking, a nonlinear Markov evolution is just a dynamical system generated by a measure-valued ordinary differential equation (ODE) with the specific feature of preserving positivity. This feature distinguishes it from a general Banach space valued ODE and yields a natural link with probability theory, both in interpreting results and in the tools of analysis. Technical complications for the sensitivity analysis, again compared with the standard theory of vector-valued ODE, lie in the specific unboundedness of generators that causes the derivatives of the solutions to nonlinear equations (with respect to parameters or initial conditions) to live in other spaces, than the evolution itself. From the probabilistic point of view, the first derivative with respect to initial data (specified by the linearized evolution around a path of nonlinear dynamics) describes the interacting particle approximation to this nonlinear dynamics (which, in turn, serves as the dynamic law of large numbers to this approximating Markov system of interacting particles), and the second derivative describes the limit of fluctuations of the evolution of particle systems around its law of large numbers (probabilistically the dynamic central limit theorem).

More precise definition is as follows. Let $\tilde{\mathcal{M}}(X)$ be a dense subset of the space $\mathcal{M}(X)$ of finite (positive Borel) measures on a polish (complete separable metric) space X (considered in its weak topology). By a nonlinear *sub-Markov* (resp. *Markov*) *propagator* in $\tilde{\mathcal{M}}(X)$ we shall mean any propagator $V^{t,r}$ of possibly nonlinear transformations of $\tilde{\mathcal{M}}(X)$ that do not increase (resp. preserve) the norm. If $V^{t,r}$ depend only on the difference $t-r$ and hence specify a semigroup, this semigroup is called nonlinear (or generalized) *sub-Markov* or *Markov* respectively.

The usual, linear, Markov propagators or semigroups correspond to the case when all the transformations are linear contractions in the whole space $\mathcal{M}(X)$. In probability theory these propagators describe the evolution of averages of Markov processes, i.e. processes whose evolution after any given time t depends on the past $X_{\leq t}$ only via the present position X_t. Loosely speaking, to any nonlinear Markov propagator there corresponds a process whose behavior after any time t depends on the past $X_{\leq t}$ via the position X_t of the process and its distribution at t.

More precisely, consider the nonlinear equation in the weak form

$$\frac{d}{dt}(g, \mu_t) = (A[\mu_t]g, \mu_t), \quad g \in C(X), \tag{3}$$

with a certain family of operators $A[\mu]$ in $C(X)$ depending on μ as a parameter and such that each $A[\mu]$ specifies a uniquely defined Markov process (say, via solution to the corresponding martingale problem, or by generating a Feller semigroup).

Suppose that the Cauchy problem for equation (3) is well posed and specifies the weakly continuous Markov semigroup T_t in $\mathcal{M}(X)$. Suppose also that for any weakly continuous curve $\mu_t \in \mathcal{P}(X)$ (the set of probability measures on X) the solutions to the Cauchy problem of the equation

$$\frac{d}{dt}(g, \nu_t) = (A[\mu_t]g, \nu_t) \tag{4}$$

define a weakly continuous propagator $V^{t,r}[\mu_.]$, $r \leq t$, of linear transformations in $\mathcal{M}(X)$ and hence a Markov process in X, with transition probabilities $p_{r,t}^{[\mu_.]}(x, dy)$. Then to any $\mu \in \mathcal{P}(X)$ there corresponds a (usual linear, but time non-homogeneous) Markov process X_t^ν in X (ν stands for an initial distribution) such that its distributions ν_t solve equation (4) with the initial condition ν. We call the family of processes X_t^μ a *nonlinear Markov process*. When each $A[\mu]$ generates a Feller semigroup and T_t acts on the whole $\mathcal{M}(X)$ (and not only on its dense subspace), the corresponding process can be also called *nonlinear Feller*. Allowing for the evolution on subsets $\tilde{\mathcal{M}}(X)$ is however crucial, as it often occurs in applications, say for the Smoluchovski or Boltzmann equation with unbounded rates.

Thus a nonlinear Markov process is a semigroup of the transformations of distributions such that to each trajectory is attached a "tangent" Markov process with the same marginal distributions. The structure of these tangent processes is not intrinsic to the semigroup, but can be specified by choosing a stochastic representation for the generator, that is of the r.h.s. of (4).

The theoretical issues that we mentioned above concerned the well-posedness of equations of type (3) and its sensitivity to various parameters and were developed in full in [3], [4], [7].

The development was carried out on the level of generality needed for applications to many agent and/or control systems dealt in Year 3.

2.3 Findings for year 3: main objectives

The last three indicative directions for further possible directions (configuration space of non-trivial geometry, controlled nonlinear quantum dynamic semigroups and full infinite-dimensional measure valued control Markov processes and games) were touched upon in book [1], Sec. 11.3, 11.4 and in book [2], Chapter 6, but were mainly left to the future research.

Our main work was around two mainstreams of competitive control problems for nonlinear Markov processes:

1) Each agent has individual payoff. This leads to mean-field games initiated for the case of underlying diffusion process by P. Caines, R. Malhame, M. Huang. Here we developed the theory for an arbitrary underlying nonlinear Markov process, see [6] and

2) Individuals fulfil the objectives of competitive leaders (generals with armies, engineers with robot-swamps, etc), where the main completed results so far are presented in [8]. This paper summarized many ideas of this project and, at the same time, opened the road for several further directions of research that were not thought about at the initial stage of the project.

Let us outline the theory for both cases in more detail.

2.4 Findings for year 3: mean-field games

Mean-field game methodology aims at describing control processes with large number N of participants by studying the limit $N \to \infty$ when the contribution of each member becomes negligible and their interaction is performed via certain mean-field characteristics, which can be expressed in terms of empirical measures. A characteristic feature of the MFG analysis is the study of a coupled system of a backward equation on functions (Hamilton-Jacobi-Bellman equation) and a forward equation on probability laws (Kolmogorov equation). We showed that the machinery of nonlinear Markov processes could serve as a natural tool for studying mean-field games with the general underlying Markov dynamics of agents (not only diffusions). More specifically, the main consistency equation of MFG can be looked at as a coupling of a nonlinear Markov process with certain controlled dynamics. Using this link we develop the MFG methodology for a wide class of underlying Markov dynamics including in particular stable and stable-like processes, as well as their various modifications like tempered stable-like process or their mixtures with diffusions.

Moreover, our abstract approach yields essential improvements even for underlying processes being diffusions. In particular, it includes the case of diffusions coefficients (not only drifts) depending on empirical measures, it allows us to get rid of the assumption of small coupling (or composite gain), to prove the crucial sensitivity estimates (to derive the regularity of HJB equations from the regularity of the Hamiltonian functions), and finally to get a full prove of convergence rate of order $1/N$.

Let us explain now the main ideas, objectives and strategy of our analysis. Suppose a position of an agent is described by a point in a locally compact separable metric space $\mathcal{X}$. A position of N agents is then given by a point in the power $\mathcal{X}^N = \mathcal{X} \times \cdots \times \mathcal{X}$ (N times). Hence the natural state space for describing the variable (but not vanishing) number of players is the union $\hat{\mathcal{X}} = \cup_{j=1}^{\infty} \mathcal{X}^j$. We denote by $C_{\text{sym}}(\mathcal{X}^N)$ the Banach spaces of symmetric (with respect to permutation of all arguments) bounded continuous functions on $\mathcal{X}^N$ and by $C_{\text{sym}}(\hat{\mathcal{X}})$ the corresponding space of functions on the full space $\hat{\mathcal{X}}$. We denote the elements of $\hat{\mathcal{X}}$ by bold letters, say $\mathbf{x}$, $\mathbf{y}$.

Reducing the set of observables to $C_{\text{sym}}(\hat{\mathcal{X}})$ means effectively that our state space is not $\hat{\mathcal{X}}$ (or $\mathcal{X}^N$ in case of a fixed number of particles) but rather the quotient space $S\hat{\mathcal{X}}$ (or $S\mathcal{X}^N$ resp.) obtained with respect to the action of the group of permutations, which allows the identifications $C_{\text{sym}}(\hat{\mathcal{X}}) = C(S\hat{\mathcal{X}})$ and $C_{\text{sym}}(\mathcal{X}^N) = C(S\mathcal{X}^N)$. Clearly $S\hat{\mathcal{X}}$ can be identified with the set of all finite collections of points from $\mathcal{X}$, the order being irrelevant.

A key role in the theory of measure-valued limits of interacting particle systems is played by the inclusion $S\hat{\mathcal{X}}$ to $\mathcal{P}(\mathcal{X})$ (the set of probability laws on $\mathcal{X}$) given by

$$\mathbf{x} = (x_1, ..., x_N) \mapsto \frac{1}{N}(\delta_{x_1} + \cdots + \delta_{x_N}) = \frac{1}{N}\delta_{\mathbf{x}}, \tag{5}$$

which defines a bijection between $S\mathcal{X}^N$ and the subset $\mathcal{P}_\delta^N(\mathcal{X})$ (of normalized sums of Dirac's masses) of $\mathcal{P}(\mathcal{X})$. This bijection extends to the bijection of $S\hat{\mathcal{X}}$ to

$$\mathcal{P}_\delta(\mathcal{X}) := \cup_{N=1}^{\infty} \mathcal{P}_\delta^N(\mathcal{X}) \subset \mathcal{P}(\mathcal{X}),$$

that can be used to equip $S\hat{\mathcal{X}}$ with the structure of a metric space by pulling back any distance on $\mathcal{P}(\mathcal{X})$ that is compatible with its weak topology.

Let $\{A[t, \mu, u]\}$ be a family of generators of Feller processes in $\mathcal{X}$, where $t \geq 0$, $\mu \in \mathcal{P}(\mathcal{X})$ and $u \in \mathbf{U}$ (a metric space interpreted as a set of admissible controls). Assume also that a mapping $\gamma : \mathbf{R}^+ \times \mathcal{X} \to \mathbf{U}$ is given. For any N, let us define the following (time-dependent) family of operators (pre-generators) on $C_{\text{sym}}(\mathcal{X}^N)$ describing N mean-field interacting agents:

$$\widehat{A}_t^N[\gamma] f(\mathbf{x}) = \widehat{A}_t^N[\gamma] f(x_1, \cdots, x_N) := \sum_{i=1}^N A^i[t, \mu, u_i] f(x_1, \cdots, x_N), \tag{6}$$

where

$$\mu = \frac{1}{N}\sum_{i=1}^N \delta_{x_i} = \frac{1}{N}\delta_{\mathbf{x}}$$

is the empirical distribution of agents, $u^i = \gamma(t, x_i)$ and $A^i[t, \mu, u_i] f$ means the action of the operator $A[t, \mu, u_i]$ on the ith variable of the function f. Let us assume that the family $\widehat{A}_t^N[\gamma]$ generates a Markov process $X^N = \{X^N(t) = (X_1^N(t), \ldots, X_N^N(t) : t \geq 0)\}$ on $\mathcal{X}^N$ for any N. We shall refer to it as a *controlled (via control γ) process of N mean-field interacting agents.*

In the terminology of statistical mechanics the operator $\hat{A}_t[\gamma]$ (considered for all N, i.e. lifted naturally to the whole space $C_{\text{sym}}(\hat{\mathcal{X}})$) should be called the second quantization of $A[t, \mu, u]$.

Using mapping (5), we can transfer our process of N mean-field interacting agents from $S\mathcal{X}^N$ to $\mathcal{P}_\delta^N(\mathcal{X})$. This leads to the following operator on $C(\mathcal{P}_\delta^N(\mathcal{X}))$:

$$\widehat{A}_t^N[\gamma] F(\delta_{\mathbf{x}}/N) = \widehat{A}_t^N[\gamma] f(\mathbf{x}) = \sum_{i=1}^N A^i[t, \mu, u_i] f(x_1, \cdots, x_N), \tag{7}$$

where $f(\mathbf{x}) = F(\delta_\mathbf{x}/N)$ and $\mathbf{x} = (x_1, \cdots, x_N)$. Let us calculate the action of this operator on linear functionals F, that is on the functionals of the form

$$F^g(\mu) = (g, \mu) = \int g(x)\mu(dx) \tag{8}$$

for a $g \in C(\mathcal{X})$. Denoting $g^\oplus(\mathbf{x}) = \sum_{i=1}^N g(x_i)$ for $\mathbf{x} = (x_1, \cdots x_N)$ we get

$$\begin{aligned} \widehat{A}_t^N[\gamma] F^g(\delta_\mathbf{x}/N) &= \frac{1}{N}\left(\widehat{A}_t^N[\gamma] g^\oplus\right)(x_1, \cdots, x_N) \\ &= \frac{1}{N}\sum_{i=1}^N \left(A[t, \delta_\mathbf{x}/N, \gamma(t, x_i)]g\right)(x_i) = \left(A[t, \delta_\mathbf{x}/N, \gamma(t, .)]g, \delta_\mathbf{x}/N\right). \end{aligned} \tag{9}$$

Hence, if $\mu_t^N = \delta_\mathbf{x}/N \to \mu_t \in \mathcal{P}(\mathcal{X})$ as $N \to \infty$, we have

$$\widehat{A}_t^N[\gamma] F^g(\delta_\mathbf{x}/N) \to \left(A[t, \mu_t^N, \gamma(t, .)]g, \mu_t^N\right), \quad \text{as} \quad N \to \infty,$$

so that the evolution equation

$$\dot{F}_t = \widehat{A}_t^N[\gamma] F_t \tag{10}$$

of our controlled process of N mean-field interacting agents, for the linear functionals of the form $F_t^g(\mu) = (g, \mu_t(\mu))$ turns to the equation

$$\frac{d}{dt}(g, \mu_t) = (A[t, \mu_t, \gamma(t, .)]g, \mu_t), \quad \mu_0 = \mu. \tag{11}$$

We call this equation the *general kinetic equation* in weak form. It should hold for g from a suitable class of test functions. This limiting procedure will be discussed in detail later on.

Let us explain how the mapping γ pops in from individual controls. Assume that the objective of each agent is to maximize (over a suitable class of controls $\{u.\}$) the payoff

$$\mathbf{E}\left[\int_t^T J(s, X_i^N(s), \mu_s^N, u_s)\, ds + V^T(X_i^N(T))\right],$$

consisting of running and final components, where the functions $J : \mathbf{R}^+ \times \mathcal{X} \times \mathcal{P}(\mathcal{X}) \times \mathbf{U} \to \mathbf{R}$ and $V^T : \mathcal{X} \to \mathbf{R}$, and the final time T are given, and where $\{\mu.\}$ is the family of the empirical measures of the whole process

$$\mu_s^N = \frac{1}{N}(\delta_{X_1^N(s)} + \cdots + \delta_{X_N^N(s)}), \quad t \le s \le T.$$

By dynamic programming (and assuming appropriate regularity), if the dynamics of empirical measures μ_s is given, the optimal payoff

$$V_N(t, x) = \sup_{u.} \mathbf{E}\left[\int_t^T J(s, X(s), \mu_s^N, u_s)\, ds + V^T(X(T))\right]$$

of an agent starting at x at time t should satisfy the HJB equation

$$\frac{\partial V_N(t, x)}{\partial t} + \max_u \left(J(t, x, \mu_t^N, u) + A[t, \mu_t^N, u]V_N(t, x)\right) = 0 \tag{12}$$

with the terminal condition $V_N(T,.) = V^T(\cdot)$. If $\mu_t^N \to \mu_t \in \mathcal{P}(\mathcal{X})$ as $N \to \infty$, then it is reasonable to expect that the solution of (12) converges to the solution of the equation

$$\frac{\partial V(t,x)}{\partial t} + \max_u \big(J(t,x,\mu_t,u) + A[t,\mu_t,u]V(t,x)\big) = 0. \tag{13}$$

Assume HJB equation (13) is well posed and the max is achieved at one point only. Let us denote this point of maximum by $u = \Gamma(t,x,\{\mu_{\geq t}\})$.

Thus, if each agent chooses the control via HJB (13), given an empirical measure $\hat{\mu}$, i.e. with

$$\gamma(t,x) = \Gamma(t,x,\{\hat{\mu}_{\geq t}\}), \tag{14}$$

this γ specifies a nonlinear Markov evolution $\{\mu_t\}_{t\geq 0}$ via kinetic equation (11). The corresponding *MFG consistency (or fixed point) condition* $\{\hat{\mu}_.\} = \{\mu_.\}$ leads to the equation

$$\frac{d}{dt}(g,\mu_t) = (A[t,\mu_t,\Gamma(t,.,\{\mu_{\geq t}\})]g,\mu_t), \tag{15}$$

which expresses the *coupling* of the nonlinear Markov process specified by (15) and the optimal control problem specified by HJB (13). It is now reasonable to expect that if the number of agents N tends to infinity in such a way that the limiting evolution is well defined and satisfies the limiting equation (15) with Γ chosen via the solution of the above HJB equation, then the control γ and the corresponding payoffs represent the ϵ-Nash equilibrium for the controlled system of N agents, with $\epsilon \to 0$, as $N \to \infty$. This statement (or conjecture) represents the essence of the *MFG methodology*.

Under certain assumptions on the family $A[t,\mu,u]$, we justify this claim by carrying out the following tasks:

T1) Proving the existence of solutions to the Cauchy problem for coupled kinetic equations (15) within an appropriate class of feedback Γ and the well-posedness for the uncoupled equations (11). Notice that we are not claiming uniqueness for (15). It is difficult to expect this, as in general Nash equilibria are not unique. At the same time, it seems to be an important open problem to better understand this non-uniqueness by describing and characterizing specific classes of solutions. On the other hand, well-posedness for the uncoupled equations (11) is crucial for further analysis.

T2) Proving the well-posedness of the Cauchy problem for the (backward) HJB equation (13), for an arbitrary flow $\{\mu.\}$ in some class of regularity, yielding the feedback function Γ in the class required by T1). This should include some sensitivity analysis of Γ with respect to the functional parameter $\{\mu_.\}$, which will be needed to show that approximating the limiting MFG distribution $\{\mu.\}$ by approximate N-particle empirical measures yields also an approximate optimal control. To perform this task, we shall assume here additionally that the operators $A[t,\mu,u]$ in (13) can be decomposed into the sum of a controlled 1st order term and a term that does not depend on control and generates a propagator with certain smoothing properties. This simplifying assumption allows to work out the theory with classical (or at least mild) solutions of HJB equations. Without this assumption, one would have to face additional technical complications related to viscosity solutions.

T3) Showing the convergence of the N-particle approximations, given by generators (9) to the limiting evolution (11), i.e. the *dynamic laws of large numbers* (LLN), for a class of controls γ

arising from (14) with a fixed $\{\hat{\mu}.\}$, where Γ is from the class required for the validity of T1) and T2). Here one can use either more probabilistic compactness and tightness (on Skorokhod paths spaces) approach, or a more analytic method via semigroups of linear operators on continuous functionals of measures. We use the second method, as it yields more precise convergence rates. For the analysis of the convergence of the corresponding semigroups the crucial ingredient is the analysis of smoothness (*sensitivity*) of the solutions to kinetic equations (11) with respect to initial data. The rates of convergence in LLN imply directly the corresponding rather precise estimates for the so-called *propagation of chaos* property of interacting particles.

T4) Finally, combining T2) and T3), one has to show that thus obtained strategic profile (14) with $\{\hat{\mu}.\} = \{\mu.\}$ represents an ϵ-*equilibrium* for N agents system with $\epsilon \to 0$, as $N \to \infty$. Actually we going to prove this with $\epsilon = 1/N$ using the method of *tagged particles* in our control setting.

This program is carried out under rather general assumptions in the extensive preprint [6].

Let us specify our model a bit further.

Of particular interest are the models with the one-particle space $\mathcal{X}$ having a spatial and a discrete components, the latter interpreted as a type of an agent. Thus let $\mathcal{X} = \mathbf{R}^d \times \mathcal{K}$, where $\mathcal{K}$ is either a finite or denumerable set. In this case, functions from $C(\mathcal{X})$ can be represented by sequences $f = (f_i)_{i \in \mathcal{K}}$ with each $f_i \in C(\mathbf{R}^d)$, the probability laws on $\mathcal{X}$ are similarly given by the sequences $\mu = (\mu_i)_{i \in \mathcal{K}}$ of positive measures on $\mathbf{R}^d$ with the masses totting up to one.

The operators A in $C(\mathcal{X})$ are specified by operator-valued matrices $\{A_{ij}\}$, $i, j \in \mathcal{K}$, with A_{ij} being an operator in $C(\mathbf{R}^d)$, so that $(Af)_i = \sum_{j \in \mathcal{K}} A_{ij} f_j$. It is not difficult to show that for such a matrix A to define a conditionally positive conservative operator in $C(\mathcal{X})$ (in particular, a generator of a Feller process) it is necessary that A_{ij} for $i \neq j$ are integral operators

$$(A_{ij} f)(z) = \int_{\mathbf{R}^d} (f_j(y) - f(z)) \nu_{ij}(z, dy)$$

with a bounded (for each z) measure $\nu_{ij}(z, dy)$, and the diagonal terms are given by the *Lévy-Khintchin* type operators ($i \in \mathcal{K}$):

$$A_{ii} f(z) = \frac{1}{2}(G_i(z)\nabla, \nabla) f(z) + (b_i(z), \nabla f(z))$$

$$+ \int_{\mathbf{R}^d} (f(z+y) - f(z) - (\nabla f(z), y) \mathbf{1}_{B_1}(y)) \nu_i(z, dy), \tag{16}$$

with $G_i(z)$ being a symmetric non-negative matrix, $\nu_i(z, .)$ being a Lévy measure on $\mathbf{R}^d$, i.e.

$$\int_{\mathbf{R}^d} \min(1, |y|^2) \nu_i(z, dy) < \infty, \quad \nu(\{0\}) = 0, \tag{17}$$

depending measurably on z, and where $\mathbf{1}_{B_1}$ denotes, as usual, the indicator function of the unit ball in $\mathbf{R}^d$.

Operators A_{ij} with $i \neq j$ describe the *mutation (migration)* between the types. If mutations are not allowed, A will be given by a diagonal matrix with the diagonal terms $A_i = A_{ii}$ of type (16).

Let us assume additionally that each agent can control only its drift, that is the diagonal generators have the form

$$A_i[t, \mu, u] f(z) = (h_i(t, z, \mu, u), \nabla f(z)) + L_i[t, \mu] f(z), \quad i = 1, \cdots, K, \tag{18}$$

with L_i of form (16), i.e.

$$
\begin{aligned}
L_i[t,\mu]f(z) &= \frac{1}{2}(G_i(t,z,\mu)\nabla,\nabla)f(z) + (b_i(t,z,\mu),\nabla f(z)) \\
&+ \int_{\mathbf{R}^d}(f(z+y) - f(z) - (\nabla f(z), y)\mathbf{1}_{B_1}(y))\nu_i(t,z,\mu,dy)
\end{aligned}
\tag{19}
$$

with the coefficients G_i, b_i, ν_i depending on $t \in \mathbf{R}^+$ and $\mu = (\mu_1, \cdots, \mu_K) \in \mathcal{P}(\mathcal{X})$ as parameters.

If, for a given (probability) measure flow $\{\mu_t\}_{t\in[0,T]}$, the operators $L[t,\mu_t] = (L_1,\cdots,L_K)[t,\mu_t]$ generate a Markov process $\{R_t[\mu_t]\}_{t\in[0,T]} = \{(R_t^1[\mu_t],\cdots,R_t^K[\mu_t])\}_{t\in[0,T]}$, one can write a stochastic differential equation (SDE) corresponding to the generator given in (18) as

$$
dX_t^i = h_i(t, X_t^i, \mu_t, u_t^i)\,dt + dR_t^i[\mu_t], \quad i = 1, \cdots, K.
$$

If μ_t are required to coincide with the laws of X_t^i, for all $t \in [0,T]$, these equations take the form of *SDEs driven by nonlinear Lévy noises*, developed in [1], [3], [7].

The initial work on the mean field games, done by Lions et al. and Caines et al., dealt with the processes $R_t[\mu]$ being Brownian Motions without dependence on μ. In our framework, this underlying process is extended to an arbitrary Markov process with a generator (19) depending on μ.

In the main kinetic equation (11), we shall then have $(g,\mu_t) = \sum_{i=1}^K (g_i, \mu_{i,t})$ and

$$
A[t,\mu_t,\gamma(t,.)]g = \{A_i[t,\mu_t,\gamma(t,.)]g_i\}_{i=1}^K
$$

with

$$
A_i[t,\mu,\gamma(t,.)]g_i(z) = (h_i(t,z,\mu,\gamma(t,.)), \nabla g_i(z)) + L_i[t,\mu]g_i(z). \tag{20}
$$

HJB equation (13) now decomposes into a collection of HJB equations for each class of agents, written as

$$
\frac{\partial V^i(t,x)}{\partial t} + H_t^i(x, \nabla V^i(x), \mu_t) + L_i[t,\mu_t]V^i(t,x) = 0 \tag{21}
$$

where

$$
H_t^i(x,p,\mu_t) := \max_{u\in\mathbf{U}}\{h_i(t,x,\mu_t,u)p + J_i(t,x,\mu_t,u)\}. \tag{22}
$$

We have assumed the resulting feedback control is unique (i.e. argmax in (22) is unique). The basic example of such situation is given by H_∞-optimal control problems, where For each i, the running cost function J_i is quadratic in u, i.e.

$$
J_i(t,x,\mu,u) = \alpha_i(t,x,\mu) - \theta_i(t,x,\mu)u^2
$$

and the drift coefficient h_i is linear in u, i.e.

$$
h_i(t,x,\mu,u) = \beta_i(t,x,\mu)u,
$$

where the functions $\alpha_i, \beta_i, \theta_i : [0,T] \times \mathbf{R}^d \times \mathcal{P}(\mathbf{R}^d) \to \mathbf{R}$ and $\theta_i(t,x,\mu) > 0$ for any (t,x,μ). Thus the explicit formula of the unique point of maximum becomes available:

$$
u = \frac{\beta}{2\theta}(t,x,\mu)p,
$$

and the HJB equation (21) rewrites as

$$\frac{\partial V^i(t,x)}{\partial t} + \frac{\beta_i^2}{4\theta_i}(t,x,\mu)(\nabla V^i)^2(t,x) + \alpha_i(t,x,\mu) + L_i[t,\mu_t]V^i(t,x) = 0$$

which is a generalized backward Burger's equation. Another natural example is the situation, where, for each i, $h_i(t,x,\mu,u) = u$ and $J_i(t,x,\mu,u)$ is a strictly concave smooth function of u. Then H_t^i is the Legendre transform of $-J$ as a function of u, the unique point of maximum in (22) is therefore $u = \partial H_t^i/\partial p$ and the kinetic equation (15) takes the form

$$\frac{d}{dt}(g,\mu_t) = \sum_{k=1}^{K}(L_i(t,\mu_t)g_i + \frac{\partial H_t^i}{\partial p}(x,p,\mu_t)|_{p=\nabla V^i(x)}\nabla g_i, \mu_{i,t}). \tag{23}$$

2.5 Findings for year 3: centralized control

Let us turn to the discussion of centralized controls. We consider here only the case of finite initial state space (for which the theory is fully developed so far), when the corresponding space of measures becomes a finite-dimensional Euclidean space (more precisely its positive orthant $\mathbf{R}_+^d$), so that the limiting measure-valued evolution becomes a deterministic control process or a differential game in $\mathbf{R}_+^d$. Let us show how the identification of deterministic limit is carried and formulate the main results on convergence referring for full proofs to [8]. we shall assume that there is a fixed number of players $\{1,\cdots,K\}$ each controlling a stochastic system consisting of a large number $N_1,\cdots,N_K \to \infty$ components respectively. These can be generals controlling armies, engineers controlling robot swamps, large banks managers controlling subsidiaries, etc. The components can interact between themselves and with agents of other groups. The limit $N_1,\cdots,N_k \to \infty$ will be described by a differential game in R_+^K.

Recall the standard notation $C^k(\Omega)$, $k \in \mathbf{N}$, for the Banach space of k times continuously differentiable functions in the interior of $\Omega \subset \mathbf{R}^d$ with f and all its derivatives up to and including order k having continuous and bounded extension to Ω, equipped with norm $\|f\|_{C^k(\Omega)}$ which is the sum of the sup-norms of f and all its derivatives up to and including order k. For $\alpha \in (0,1]$, we denote by $C^{k,\alpha}(\Omega)$ the subspace of $C^k(\Omega)$ consisting of functions, whose kth order derivatives are Hölder continuous of index α. The Banach norm on this space is defined as the sum of the norm in $C^k(\Omega)$ plus the minimal Hölder constant.

Law of large numbers for interacting Markov chains.

Let us first recall the basic setting of mean-field interacting particle systems with a finite number of types. Suppose our initial state space is a finite set $\{1,...,d\}$, which can be interpreted as the types of particles (say, possible opinions of individuals on a certain subject, or the levels of fitness in a military unit, or the types of robots in a robot swamp). Let $\{Q(t,x)\} = \{(Q_{ij})(t,x)\}$ be a family of $d \times d$ square Q-matrices or Kolmogorov matrices (i.e. non-diagonal elements of these matrices are non-negative and the elements of each row sum up to one) depending continuously on a vector x from the closed simplex

$$\Sigma_d = \{x = (x_1,...,x_d) \in \mathbf{R}_+^d : \sum_{j=1}^{d} x_j = 1\},$$

and piecewise continuously on time $t \geq 0$. For any x, the family $\{Q(., x)\}$ specifies a Markov chain on the state space $\{1, ..., d\}$ with the generator

$$(Q(t,x)f)_n = \sum_{m \neq n} Q_{nm}(t,x)(f_m - f_n), \quad f = (f_1, \cdots, f_d),$$

and with the intensity of jumps being

$$|Q_{ii}(t,x)| = -Q_{ii}(t,x) = \sum_{j \neq i} Q_{ij}(t,x).$$

In other words, the transition matrices $P(s,t,x) = (P_{ij}(s,t,x))_{i,j=1}^d$ of this chain satisfies the Kolmogorov forward equations

$$\frac{d}{dt} P_{ij}(s,t,x) = \sum_{l=1}^{d} Q_{lj}(t,x) P_{il}(s,t,x), \quad s \leq t.$$

Suppose we have a large number of particles distributed arbitrary among the types $\{1, ..., d\}$. More precisely our state space S is $\mathbf{Z}_+^d$, the set of sequences of d non-negative integers $N = (n_1, ..., n_d)$, where each n_i specifies the number of particles in the state i. Let $|N|$ denote the total number of particles in state N: $|N| = n_1 + ... + n_d$. For $i \neq j$ and a state N with $n_i > 0$ denote by N^{ij} the state obtained from N by removing one particle of type i and adding a particle of type j, that is n_i and n_j are changed to $n_i - 1$ and $n_j + 1$ respectively. The mean-field interacting particle system specified by the family $\{Q\}$ is defined as the Markov process on S specified by the generator

$$L_t f(N) = \sum_{i,j=1}^{d} n_i Q_{ij}(t, N/|N|)[f(N^{ij}) - f(N)]. \tag{24}$$

Probabilistic description of this process is as follows. Starting from any time and current state N one attaches to each particle a $|Q_{ii}|(N/|N|)$-exponential random waiting time (where i is the type of this particle). If the shortest of the waiting times τ turns out to be attached to a particle of type i, this particle jumps to a state j according to the distribution $(Q_{ij}/|Q_{ii}|)(N/|N|)$. Briefly, with this distribution and at rate $|Q_{ii}|(N/|N|)$, any particle of type i can turn (migrate) to a type j. After any such transition the process starts again from the new state N^{ij}. Notice that since the number of particles $|N|$ is preserved by any jump, this process is in fact a Markov chain with a finite state space.

Remark 1 *Yet another way of describing the chain generated by L_t is via the forward Kolmogorov (or master) equation for its transition probabilities $P_{MN}(s,t)$:*

$$\frac{d}{dt} P_{MN}(s,t) = \sum_{i,j=1}^{d} (n_i + 1) Q_{ij}(t, \frac{N^{ji}}{|N|}) P_{MN^{ij}}(s,t) - \sum_{i,j=1}^{d} n_i Q_{ij}(t, \frac{N^{ij}}{|N|}) P_{MN}(s,t), \quad s \leq t.$$

To shorten the formulas, we shall denote the inverse number of particles by h, that is $h = 1/|N|$. Normalizing the states to $N/|N| \in \Sigma_d^h$, where Σ_d^h is a subset of Σ_d with coordinates proportional to h, leads to the generator of the form

$$L_t^h f(N/|N|) = \sum_{i=1}^d \sum_{j=1}^d \frac{n_i}{|N|} |N| Q_{ij}(t, N/|N|)[f(N^{ij}/|N|) - f(N/|N|)], \tag{25}$$

or equivalently

$$L_t^h f(x) = \sum_{i=1}^d \sum_{j=1}^d x_i Q_{ij}(t,x) \frac{1}{h}[f(x - he_i + he_j) - f(x)], \quad x \in h\mathbf{Z}_+^d, \tag{26}$$

where $e_1, ..., e_d$ denotes the standard basis in $\mathbf{R}^d$. With some abuse of notation, let us denote by $hN^{t,h}$ the corresponding Markov chain. The transition operators of this chain will be denoted by $\Psi_{s,t}^h$:

$$\Psi_{s,t}^h f(hN) = \mathbf{E}_{s,hN} f(hN(t,h)), \quad s \le t, \tag{27}$$

where $\mathbf{E}_{s,x}$ denotes the expectation of the chain started at x at time s. These operators are known to form a propagator, i.e. they satisfy the chain rule (or Chapman-Kolmogorov equation)

$$\Psi_{s,t}^h \Psi_{t,r}^h = \Psi_{s,t}^h, \quad s \le t \le r.$$

We shall be interested in the asymptotic behavior of these chains as $h \to 0$. To this end, let us observe that, for $f \in C^1(\Sigma_d)$,

$$\lim_{|N|\to\infty,\, N/|N|\to x} |N|[f(N^{ij}/|N|) - f(N/|N|)] = \frac{\partial f}{\partial x_j}(x) - \frac{\partial f}{\partial x_i}(x),$$

so that

$$\lim_{|N|\to\infty,\, N/|N|\to x} L_t^h f(N/|N|) = \Lambda_t f(x),$$

where

$$\Lambda_t f(x) = \sum_{i=1}^d \sum_{j\neq i} x_i Q_{ij}(t,x)[\frac{\partial f}{\partial x_j} - \frac{\partial f}{\partial x_i}](x) = \sum_{k=1}^d \sum_{i\neq k} [x_i Q_{ik}(t,x) - x_k Q_{ki}(t,x)] \frac{\partial f}{\partial x_k}(x). \tag{28}$$

The limiting operator $\Lambda_t f$ is a first-order PDO with characteristics solving the equation

$$\dot{x}_k = \sum_{i\neq k} [x_i Q_{ik}(t,x) - x_k Q_{ki}(t,x)] = \sum_{i=1}^d x_i Q_{ik}(t,x), \quad k = 1, ..., d, \tag{29}$$

called the *kinetic equations* for the process of interaction described above. The characteristics specify the dynamics of the deterministic time-nonhomogeneous Markov Feller process in Σ_d defined via the generator Λ_t. The corresponding transition operators act on $C(\Sigma_d)$ as

$$\Phi_{s,t} f(x) = f(X_{s,x}(t)), \quad s \le t, \tag{30}$$

where $X_{s,x}(t)$ is the solution to (29) with the initial condition x at time s. These operators form a Feller propagator (i.e. $\Phi_{s,t}$ depend strongly continuous on s, t and satisfy the chain rule $\Phi_{s,t}\Phi_{t,r} = \Phi_{s,r}$, $s \le t \le r$). Of course in case of Q that do not depend on time t explicitly, $\Phi_{s,t}$ depend only on the difference $t - s$ and the operators $\Phi_t = \Phi_{0,t}$ form a Feller semigroup.

Remark 2 *It is easy to see that if $x_k \neq 0$, then $(X_{s,x}(t))_k \neq 0$ for any $t \geq s$. Hence the boundary of Σ_d is not attainable for this semigroup, but, depending on Q, it can be glueing or not. For instance, if all elements of Q never vanish, then the points $X_{s,x}(t)$ never belong to the boundary of Σ_d for $t > s$, even if the initial point x does so.*

Theorem 2.1 *(i) Let all the elements $Q_{ij}(t,.)$ belong to $C^{1,\alpha}(\Sigma)$, $\alpha \in (0,1]$, with norms uniformly bounded in t. Then, if for some $s > 0$ and $x \in \mathbf{R}^d$, the initial data hN_s converge to x in $\mathbf{R}^d$, as $h \to 0$, the Markov chains $hN(t,h)$ with the initial data hN_s (generated by L_t^h and with transitions $\Psi_{s,t}$) converge in distribution and in probability to the deterministic characteristic $X_{s,x}(t)$. For the corresponding converging propagators of transition operators the following rates of convergence hold:*

$$\sup_{0\leq s\leq t\leq T} \sup_{N\in \mathbf{Z}_+^d : |N|=1/h} \left[\Psi_{s,t}^h f(hN) - \Phi_{s,t} f(hN)\right] \leq C(T)(t-s)h^\alpha \|f\|_{C^{1,\alpha}(\Sigma_d)}, \tag{31}$$

for $f \in C^{1,\alpha}(\Sigma)$ and

$$\sup_{0\leq s\leq t\leq T} [\mathbf{E}_{s,hN} f(hN(t,h)) - f(X_{s,x}(t))] \leq C(T) \left((t-s)h^\alpha \|f\|_{C^{1,\alpha}(\Sigma_d)} + \|f\|_{C^1(\Sigma_d)}|hN - x|\right), \tag{32}$$

where $C(T)$ depends only on the supremum in t of $C^{1,\alpha}(\Sigma)$-norm of the functions $Q(t,x)$.

(ii) Assuming a weaker regularity condition, namely that $Q_{ij}(t,.)$ belong to $C^1(\Sigma)$ uniformly in t, the convergence of Markov chains $hN(t,h)$ in distribution and in probability to the deterministic characteristics still holds, but instead of (31)*, we have weaker rates in terms of the modulus of continuity w_h of ∇f and Q:*

$$\sup_{0\leq s\leq t\leq T} \sup_{N\in \mathbf{Z}_+^d : |N|=1/h} \left[\Psi_{s,t}^h f(hN) - \Phi_{s,t} f(hN)\right]$$

$$\leq C(T)(t-s)\left(w_{hC(T)}(\nabla f) + w_{hC(T)}(\nabla Q)\|f\|_{C^1(\Sigma_d)}\right), \tag{33}$$

where $C(T)$ depends on the $C^1(\Sigma)$-norm of Q. A similar modification of (32) *holds.*

Our objective is to extend this result to interacting and competitively controlled families of Markov chains.

Mean field Markov control

Turning to control dynamics, let us start with mean-field controlled Markov chains without competition. Suppose we are given a family of Q-matrices $\{Q(t,u,x)\} = \{(Q_{ij})(t,u,x),\ i,j = 1,\cdots d\}$, depending on $x \in \Sigma_d$, $t \geq 0$ and a parameter u from a metric space interpreted as control. The main assumption will be that $Q \in C^{1,\alpha}(\Sigma_d)$ as a function of x with the norm bounded uniformly in t,u, and Q depends continuously on t and u.

Any given bounded measurable curve $u(t)$, $t \in [0,T]$, defines a Markov chain on Σ_d^h with the time-dependent family of generators of type (25), that is

$$L_{t,u(t)} f\left(\frac{N}{|N|}\right) = \sum_{i,j}^d n_i Q_{ij}\left(t, u(t), \frac{N}{|N|}\right)\left[f\left(\frac{N^{ij}}{|N|}\right) - f\left(\frac{N}{|N|}\right)\right], \tag{34}$$

or equivalently

$$L^h_{t,u(t)}f(x)=\sum_{i=1}^d\sum_{j=1}^d x_iQ_{ij}(t,u(t),x)\frac{1}{h}[f(x-he_i+he_j)-f(x)]. \tag{35}$$

For simplicity (and effectively without loss of generality), we shall stick further to controls $u(.)$ from the class $C_{pc}[0,T]$ of piecewise-continuous curves (with a finite number of discontinuities).

Again for $f\in C^1(\Sigma_d)$,

$$\lim_{h=1/|N|\to 0,\, N/|N|\to x} L^h_{t,u(t)}f(N/|N|)=\Lambda_{t,u(t)}f(x),$$

where

$$\Lambda_{t,u(t)}f(x)=\sum_{k=1}^d\sum_{i\neq k}[x_iQ_{ik}(t,u(t),x)-x_kQ_{ki}(t,u(t),x)]\frac{\partial f}{\partial x_k}(x), \tag{36}$$

with the corresponding controlled characteristics governed by the equations

$$\dot{x}_k=\sum_{i\neq k}[x_iQ_{ik}(t,u(t),x)-x_kQ_{ki}(t,u(t),x)]=\sum_{i=1}^d x_iQ_{ik}(t,u(t),x),\quad k=1,...,d. \tag{37}$$

For a given $T>0$ and continuous functions J (current payoff) and V_T (terminal payoff), let $\Gamma(T,h)$ denote the problem of a centralized controller of the chain with $|N|=1/h$ particles, aiming at maximizing the payoff

$$\int_0^T J\left(s,u(s),\frac{N(s,h)}{|N|}\right)ds+V_T\left(\frac{N(T,h)}{|N|}\right). \tag{38}$$

The optimal payoff will be denoted by $V^h(t,x)$:

$$V^h(t,x)=\sup_{u(.)\in C_{pc}[t,T]}\mathbf{E}^{u(.)}_{t,x}\left[\int_t^T(J(s,u(s),hN(s,h))ds+V_T(hN(T,h))\right], \tag{39}$$

where $E^{u(.)}_{t,x}$ denotes the expectation with respect to the Markov chain on Σ^h_d generated by (34) and started at $x=hN$ at time t.

We are aiming at approximating $V^h(t,x)$ by the optimal payoff

$$V(t,x)=\sup_{u(.)\in C_{pc}[t,T]}\left[\int_t^T J(s,u(s),X_{t,x}(s))\,ds+V_T(X_{t,x}(T))\right] \tag{40}$$

for the controlled dynamics (37).

We can also obtain approximate optimal synthesis for problems $\Gamma(T,h)$ with large $|N|=1/h$, at least if regular enough synthesis is available for the limiting system. Let us recall that a function $\gamma(t,x)$ is called an optimal synthesis (or an adaptive policy) for the problem $\Gamma(T,h)$ if

$$V^h(t,x)=\mathbf{E}^{\gamma}_{t,x}\left[\int_t^T(J(s,\gamma(s,hN(s,h)),hN(s))ds+V_T(hN(T,h))\right] \tag{41}$$

for all $t \le T$ and $x \in \Sigma_d^h$, where $E_{t,x}^\gamma$ denotes the expectation with respect to the Markov chain on Σ_d^h generated by (34) with $u(t) = \gamma(t,x)$ and starting at $x = hN$ at time t. A function $\gamma(t,x)$ is called an ϵ-optimal synthesis or an ϵ-adaptive policy, if the r.h.s. of (41) differs from its l.h.s. by not more than ϵ. Similarly an optimal synthesis or an adaptive policy are defined for the limiting deterministic system.

Theorem 2.2 *(i) Assume that Q, J depend continuously on t, u and $Q, J \in C^{1,\alpha}(\Sigma_d)$, $\alpha \in (0,1]$, as functions of x, with the norms bounded uniformly in t, u, and finally $V_T \in C^{1,\alpha}(\Sigma_d)$. Then*

$$\sup_{0\le t\le T} [V^h(t,hN) - V(t,x)]$$

$$\le C(T)((T-t)h^\alpha + |hN - x|) \left(\|V_T\|_{C^{1,\alpha}(\Sigma_d)} + \sup_{s,u} \|J(t,u,.)\|_{C^{1,\alpha}(\Sigma_d)} \right), \qquad (42)$$

with $C(T)$ depending only on the bounds of the norms of Q in $C^{1,\alpha}(\Sigma_d)$. Moreover, if $u(t)$ is an ϵ-optimal control for deterministic dynamics (37), that is the payoff obtained by using $u(.)$ differs by ϵ from $V(t,x)$, then $u(.)$ is also an $(\epsilon + C(T)h^\alpha)$-optimal control for $|N| = 1/h$ particle system.

(ii) Suppose additionally that u belong to a convex subset of a Euclidean space and that $Q(t,u,x)$ depends Lipschitz continuously on u. Let $\epsilon \ge 0$, and let $\gamma(t,x)$ be a Lipschitz continuous function of x uniformly in t that represents an ϵ-optimal synthesis for the limiting deterministic control problem. Then, for any $\delta > 0$, there exists h_0 such that, for $h \le h_0$, $\gamma(t,x)$ is an $(\epsilon+\delta)$-optimal synthesis for the approximate optimal problem $\Gamma(T,h)$ on Σ_d^h.

Notice finally that by the standard dynamic programming, the optimal payoff $V(t,x)$ given by (40) represents the unique viscosity solution of the HJB-Isaacs equation

$$\frac{\partial V}{\partial t}(t,x) + \max_u \left[J(t,u,x) + \sum_{i,k=1}^d x_i Q_{ik}(u,x) \frac{\partial V}{\partial x_k}(t,x) \right] = 0, \qquad (43)$$

and the optimal payoff $V^h(t,x)$ given by (39) solves the HJB equation

$$\frac{\partial V^h}{\partial t}(t,x) + \max_u [J(t,u,x) + L_{t,u}^h V^h(t,x)] = 0. \qquad (44)$$

Thus, as a corollary of Theorem 2.2, we have proved the convergence of the solutions of the Cauchy problem for equation (44) to the viscosity solution of (43).

Two players with mean-field or binary interaction

Let us turn to a game-theoretic setting starting with a simplest model of two competing mean-field interacting Markov chains. Suppose we are given two families of Q-matrices $\{Q(t,u,x) = (Q_{ij})(u,x)\}$ and $\{P(t,v,x) = (P_{ij})(v,x)\}$, $i,j = 1, \cdots d$, depending on $x \in \Sigma_d$ and parameters u and v from two subsets U and V of Euclidean spaces. Any given bounded measurable curves $u(t), v(t)$, $t \in [0,T]$, define a Markov chain on $\Sigma_d^{1/|N|} \times \Sigma_d^{1/|M|}$, specified by the generator

$$L_{t,u(t),v(t)} f(\frac{N}{|N|}, \frac{M}{|M|}) = \sum_{i,j}^d n_i Q_{ij}(t,u(t),\frac{N}{|N|})[f\left(\frac{N^{ij}}{|N|}, \frac{M}{|M|}\right) - f\left(\frac{N}{|N|}, \frac{M}{|M|}\right)]$$

$$+\sum_{i,j}^{d} m_i P_{ij}(t,v(t),\frac{M}{|M|})[f\left(\frac{N}{|N|},\frac{M^{ij}}{M}\right)-f\left(\frac{N}{|N|},\frac{M}{|M|}\right)], \tag{45}$$

where $N=(n_1,\cdots,n_d)$, $M=(m_1,\cdots,m_d)$.

We shall assume for simplicity that $|N|=|M|=1/h$.

Then (45) rewrites as

$$L^h_{t,u(t),v(t)}f(x,y)=\sum_{i=1}^{d}\sum_{j=1}^{d} x_i Q_{ij}(t,u(t),x)\frac{1}{h}[f(x-he_i+he_j,y)-f(x,y)]$$

$$+\sum_{i=1}^{d}\sum_{j=1}^{d} y_i P_{ij}(t,v(t),y)\frac{1}{h}[f(x,y-he_i+he_j)-f(x,y)], \quad x,y\in h\mathbf{Z}^d_+. \tag{46}$$

For $f\in C^1(\Sigma_d\times\Sigma_d)$,

$$\lim_{h\to 0,\, N/|N|\to x, M/|M|\to y} L^h_{t,u(t),v(t)}f(N/|N|,M/|M|)=\Lambda_{t,u(t),v(t)}f(x,y),$$

where

$$\Lambda_{t,u(t),v(t)}f(x,y)=\sum_{k=1}^{d}\sum_{i\neq k}[x_iQ_{ik}(t,u(t),x)-x_kQ_{ki}(t,u(t),x)]\frac{\partial f}{\partial x_k}(x)$$

$$+\sum_{k=1}^{d}\sum_{i\neq k}[y_iP_{ik}(t,u(t),x)-y_kP_{ki}(t,v(t),y)]\frac{\partial f}{\partial y_k}(y). \tag{47}$$

The corresponding controlled characteristics are governed by the equations

$$\dot{x}_k=\sum_{i\neq k}[x_iQ_{ik}(t,u(t),x)-x_kQ_{ki}(t,u(t),x)]=\sum_{i=1}^{d} x_iQ_{ik}(t,u(t),x), \quad k=1,...,d, \tag{48}$$

$$\dot{y}_k=\sum_{i\neq k}[y_iP_{ik}(t,v(t),y)-y_kP_{ki}(t,v(t),y)]=\sum_{i=1}^{d} y_iP_{ik}(t,v(t),y), \quad k=1,...,d. \tag{49}$$

For a given $T>0$, let us denote by $\Gamma(T,h)$ the stochastic game with the dynamics specified by the generator (45) and with the objective of the player I (controlling Q via u) to maximize the payoff

$$\int_0^T J\left(s,u(s),v(s),\frac{N(s,h)}{|N|},\frac{M(s,h)}{|M|}\right)ds+V_T\left(\frac{N(T,h)}{|N|},\frac{M(T,h)}{|M|}\right) \tag{50}$$

for given functions J (current payoff) and V_T (terminal payoff), and with the objective of player II (controlling P via v) to minimize this payoff (zero-sum game). As previously we want to approximate it by the deterministic zero-sum differential game $\Gamma(T)$, defined by dynamics (48), (49) and the payoff of player I given by

$$\int_0^T J(s,u(s),v(s),X_{t,x}(s),Y_{t,y}(s))\,ds+V_T(X_{t,x}(T),Y_{t,y}(T)). \tag{51}$$

Recall the basic notions of the upper and lower values for a game $\Gamma(T)$. As above, we shall use controls $u(.)$ and $v(.)$ from the classes $C_{pc}([0,T];U)$ and $C_{pc}([0,T];V)$ of piecewise-continuous curves with values in U and V respectively. A progressive strategy of player I is defined as a mapping β from $C_{pc}([0,T];V)$ to $C_{pc}([0,T];U)$ such that if $v_1(.)$ and $v_2(.)$ coincide on some initial interval $[0,t]$, $t < T$, then so do $u_1 = \beta(v_1(.))$ and $u_2 = \beta(v_2(.))$. Similarly progressive strategies are defined for player II. Let us denote the sets of progressive strategies for players I and II by $S_p([0,T];U)$ and $S_p([0,T];V)$. Then the upper and the lower values for the game $\Gamma(T)$ are defined as

$$V_+(t,x,y) = \sup_{\beta \in S_p([0,T];U)} \inf_{v(.) \in C_{pc}([0,T];V)}$$

$$\left[\int_t^T J(s,(\beta(v))(s),v(s),X_{t,x}(s),Y_{t,x}(s))\,ds + V_T(X_{t,x}(T),Y_{t,x}(T))\right], \tag{52}$$

$$V_-(t,x,y) = \inf_{\beta \in S_p([0,T];V)} \sup_{u(.) \in C_{pc}([0,T];U)}$$

$$\left[\int_t^T J\left(s,u(s),(\beta(u))(s),X_{t,x}(s),Y_{t,x}(s)\right)\,ds + V_T\left(X_{t,x}(T),Y_{t,x}(T)\right)\right].$$

If the so called Isaac's condition holds, that is, for any p_k, q_k,

$$\max_u \min_v \left[J(t,u,v,x,y) + \sum_{i,k=1}^d x_i Q_{ik}(t,v,x)q_k + \sum_{i,k=1}^d y_i P_{ik}(t,v,x)p_k\right]$$

$$= \min_v \max_u \left[J(t,u,v,x,y) + \sum_{i,k=1}^d x_i Q_{ik}(t,v,x)q_k + \sum_{i,k=1}^d y_i P_{ik}(t,v,x)p_k\right], \tag{53}$$

then the upper and lower values coincide: $V_+(t,x,y) = V_-(t,x,y)$.

Similarly the upper and the lower values $V_+^h(t,x,y)$ and $V_-^h(t,x,y)$ for the stochastic game $\Gamma(T,h)$ are defined.

Theorem 2.3 *Assume that Q,P,J depend continuously on t,u and $Q,P,J,V_T \in C^{1,\alpha}(\Sigma_d)$, $\alpha \in (0,1]$, as functions of x, with the norms bounded uniformly in t,u,v. Then*

$$\sup_{0\le t\le T} [V_\pm^h(t,hN) - V_\pm(t,x)]$$

$$\le C(T)((T-t)h^\alpha + |hN-x|)\left(\|V_T\|_{C^{1,\alpha}(\Sigma_d)} + \sup_{s,u}\|J(t,u,v,.)\|_{C^{1,\alpha}(\Sigma_d)}\right), \tag{54}$$

with $C(T)$ depending only on the bounds of the norms of Q in $C^{1,\alpha}(\Sigma_d)$. Moreover, if $\beta \in S_p([0,T];U)$ and $v(.) \in C_{pc}([0,T];V)$ are ϵ-optimal for the minimax problem (52)*, then this pair is also $(\epsilon + C(T)h^\alpha)$-optimal for the corresponding stochastic game $\Gamma(T,h)$.*

As in Theorem 2.2 (ii), one can also approximate optimal (equilibrium) adaptive polices for $\Gamma(T,h)$, if regular enough (i.e. Lipschitz continuous) equilibrium adaptive policies exist for the limiting game $\Gamma(T)$. In fact, as is known from differential games, the upper value $V_+(t,x,y)$ represents the unique viscosity solution of the upper Isaac's equation

$$\frac{\partial V_+}{\partial t}(t,x,y) + \min_v \max_u \left[J(t,u,v,x,y) + \Lambda_{t,u,v} V_+(t,x,y)\right], \quad V_+(T,x,y) = V_T(x,y), \tag{55}$$

and $V_-(t,x,y)$ of the lower Isaac's equation (with min and max placed in a different order). Similar equations are satisfied by the values of stochastic games $V_\pm^h(t,x,y)$ (see e.g. [?]). Now, if V^* is a solution to the Cauchy problem (55) and there exist Lipschitz continuous functions $v^*(t,x,y)$ and $u^*(t,v,x,y)$ such that

$$u^*(t,v,x,y) \in argmax[J(t,u,v,x,y) + \Lambda_{t,u,v} V^*(t,x,y)],$$

$$v^*(t,x,y) \in argmin \max_v [J(t,u,v,x,y) + \Lambda_{t,u,v} V^*(t,x,y)],$$

then V^* is a saddle point for the differential game $\Gamma^+(T)$ giving the information advantage to maximizing player I. Analogously to Theorem 2.2 (ii), we can conclude by Theorem 2.3 that the policies $v^*(t,x,y)$ and $u^*(t,v,x,y)$ represent ϵ-equilibria for the corresponding stochastic game $\Gamma^+(T,h)$.

In a slightly different setting one can assume that changes in a competitive control process occur as a result of group interactions, and are not determined just by the overall mean field distribution. Let us discuss a simple situation with binary interaction. Assume we have two groups of d states (of objects or agents) controlled by players I and II respectively. Suppose now that any particle from a state i of the first group can interact with any particle from a state j of the second group (binary interaction) producing changes i to l and j to r with certain rates $Q_{ij}^{lr}(t,u,v)$ that may depend on controls u and v of the players. Assuming, as usual, that our particles are indistinguishable (any particle from a state is selected for interaction with equal probability), leads to the process, generated by the operators

$$L_{t,u(t),v(t)} f(N,M) = \sum_{i,j,l,r=1}^d n_i m_j Q_{ij}^{lr}(t,u(t),v(t),\frac{N}{|N|},\frac{M}{|M|})[f(N^{il},M^{jr}) - f(N,M)].$$

Again let us assume for simplicity that $|M| = |N|$ and define $h = 1/|N| = 1/|M|$. To get a reasonable scaling limit, it is necessary to scale time by factor h leading to the generators

$$L^h_{t,u(t),v(t)} f(\frac{N}{|N|},\frac{M}{|M|}) = h \sum_{i,j,l,r=1}^d n_i m_j Q_{ij}^{lr}(t,u(t),v(t),\frac{N}{|N|},\frac{M}{|M|})[f(N^{il},M^{jr}) - f(N,M)], \tag{56}$$

which, for $x = hN$, $y = hM$ and $h \to 0$, tends to

$$\Lambda_{t,u(t),v(t)} f(x,y) = \sum_{i,j,l,r=1}^d x_i y_j Q_{ij}^{lr}(t,u(t),v(t),x,y) \left[\frac{\partial f}{\partial x_l} + \frac{\partial f}{\partial y_r} - \frac{\partial f}{\partial x_i} - \frac{\partial f}{\partial y_j}\right](x,y). \tag{57}$$

The corresponding kinetic equations (characteristics of this first order partial differential operator) have the form

$$\dot{x}_k = \sum_{i,j,r=1}^d y_j \left[x_i Q_{ij}^{kr}(t,u(t),v(t)) - x_k Q_{kj}^{ir}(t,u(t),v(t))\right],$$

$$\dot{y}_k = \sum_{i,j,l=1}^{d} x_i \left[y_j Q_{ij}^{lk}(t,u(t),v(t)) - y_k Q_{ik}^{lj}(t,u(t),v(t)) \right],$$

As in the previous section, we are interested in the zero-sum stochastic game, which will again be denoted by $\Gamma(T,h)$, with the dynamics specified by generator (56) and with the objective of the player I (controlling Q via u) to maximize the payoff of the same type (50), and in an approximation of this game by the limiting deterministic zero-sum differential game $\Gamma(T)$, defined by the payoff (51) of player I.

Theorem 2.4 *Assume that Q, J depend continuously on t, u, v and $Q, J, V_T \in C^{1,\alpha}(\Sigma_d)$, $\alpha \in (0,1]$, as functions of x, with the norms bounded uniformly in t, u, v. Then the same estimate* (54) *holds for the difference of upper and lower values of limiting and approximating games.*

The theory was also partially extended to the case of K players.

3 Supported Personnel

Wei Yung (PA) Postdoctoral research assistant, University of Warwick, Coventry, UK (March 2009- August 2012)

Vassili N. Kolokoltsov (PI) Reader (from 2012 Professor), University of Warwick, Coventry, UK

4 Publications

[1] V. N. Kolokoltsov. *Nonlinear Markov processes and kinetic equations*. Cambridge Tracks in Mathematics 182, Cambridge Univ. Press, 2010. See the review of Prof. D. Applebaum in Bull. London Math. Soc. (2011) 43(6): 1245-1247.

[2] V. N. Kolokoltsov. *Markov processes, semigroups and generators*. DeGruyter Studies in Mathematics v. 38, DeGruyter, 2011.

[3] V. N. Kolokoltsov. Stochastic integrals and SDE driven by nonlinear Lévy noise. In D. Crisan (Ed.) "Stochastic Analysis in 2010", Springer 2011, p. 227-242.

[4] V. N. Kolokoltsov. Nonlinear Lévy and nonlinear Feller processes: an analytic introduction, 2011. http://arxiv.org/abs/1103.5591. To appear in De Gruyter volume "Mathematics and Life Sciences".

[5] V. N. Kolokoltsov. Game theoretic analysis of incomplete markets: emergence of probabilities, nonlinear and fractional Black-Scholes equations. http://arxiv.org/abs/1105.3053 To appear in "Risk and Decision Analysis".

[6] Vassili N. Kolokoltsov, Jiajie Li and Wei Yang. Mean Field Games and Nonlinear Markov Processes (2011). arXiv:1112.3744

[7] V. Kolokoltsov. The Lévy-Khintchine type operators with variable Lipschitz continuous coefficients generate linear or nonlinear Markov processes and semigroups. arXiv:0911.5688 (2009). Prob. Theory Related Fields 151 (2011), 95-123.

[8] V. N. Kolokoltsov. Nonlinear Markov games on a finite state space (mean-field and binary interactions) International Journal of Statistics and Probability 1:1 (2012), 77-91.

[9] Vassili Kolokoltsov and Wei Yang (2012). The turnpike theorems for Markov games. http://arxiv.org/abs/1203.6553. Published in “Dynamic Games and Applications” **2: 3** (2012), 294-312.

[10] V. N. Kolokoltsov 2012. Nonlinear diffusions and stable-like processes with coefficients depending on the median or VaR http://arxiv.org/abs/1207.5925. Submitted to 'Applied Mathematics and Optimization'.

Apart from the full publications, PI (V Kolokoltsov) and RA (Wei Yang) made several presentations on prestigious conferences and seminars, most of them resulting in electronic or on-line conference proceedings. These include

Pisa, Italy (June 2010). 16th Conference of the International Linear Algebra Society. Talk: 'Nonlinear Markov games'

Budapest (July 2010). 19th International Symposium on Mathematical Theory of Networks and Systems MITS10. Special session was organized (by PI jointly with professor McEneany) on the 'Nonlinear Markov control', that is on the theme of the grant

2010 (August 2010) Arlington VA USA 'Dynamics and control program review'. Talk: Nonlinear Markov games.

Kiev (August 2010) Humbolt Kollege 'Stochastic Analysis Workshop'. Talk: 'Nonlinear Markov processes'

SIAM bi-annual Conference on Control and its Applications, Baltimore Maryland, 25-27 July 2011; Talk: “Recent developments in Nonlinear Markov control processes”

2011 (June 14-16) Arlington VA - AFOSR Dynamics and Control Program Review. Talk “Nonlinear Markov control processes and games”.

2011 (July 6 - 8) the Royal Institute of Technology (KTH) in Stockholm, Sweden - The 16th INFORMS Applied Probability Conference. Talk “Nonlinear controlled Markov processes”

2011 (September 12-16) Vienna – ECCS'11 (European Conference on Complex Systems). Talk “Nonlinear Markov Control Processes and Games”.

2012 (22-23 May) Linne University Växjö Sweden – Guest Lectures 'Introduction to nonlinear Markov processes' and (24 - 25 May) Stochastic Analysis and Applications workshop. Talk: “Nonlinear Markov battles”.

2012 (18-22 June) Bielefeld University International Workshop ‘Qualitative Behavior of Stochastic Systems and Applications’. Talk: “Nonlinear Markov processes”.

2012 (18-22 July) Bysice, Czech Republic - XV International Symposium on Dynamic Games and Applications. Talks: “Mean field games and nonlinear Markov processes” and “Game-theoretic analysis of rainbow options in incomplete markets”

2012 (6-9 Aug) Arlington VA –Control and Dynamics Program Review. Talk “Nonlinear Markov battles”