This is a repository copy of *Mean-field interactions among robust dynamic coalitional games with transferable utilities*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/89666/

Proceedings Paper:
Bauso, D. and Basar, T. (2014) Mean-field interactions among robust dynamic coalitional games with transferable utilities. In: ISCCSP 2014 - 2014 6th International Symposium on Communications, Control and Signal Processing, Proceedings. 2014 6th International Symposium on Communications, Control, and Signal Processing (ISCCSP), 21-23 May 2014, Athens, Greece. , 487 - 490. ISBN 9781479928903

https://doi.org/10.1109/ISCCSP.2014.6877919

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

MEAN–FIELD INTERACTIONS AMONG ROBUST DYNAMIC COALITIONAL GAMES WITH TRANSFERABLE UTILITIES

Dario Bauso and Tamer Başar

DICGIM, Università di Palermo, Italy, dario.bauso@unipa.it
CSL, University of Illinois at Urbana-Champaign, US, basar1@illinois.edu

ABSTRACT

This paper considers a large number of homogeneous "small worlds" or games. Each small world involves a set of players and a corresponding set of possible coalitions, and is modeled as a dynamic game with transferable utilities (TU), where the characteristic function is a continuous-time stochastic process. Considering that a dynamic TU game can be modeled as a network control problem, the overall system appears as an assembly of a large number of networks subject to mean-field interactions. As a result of such mean-field interactions among small worlds, in each game, a central planner allocates revenues based on the extra reward that a coalition has received up to the current time and the extra reward that the same coalition has received in the other games. We obtain allocation rules that make the grand coalition stable in each game, while guaranteeing consensus on the excesses, in the spirit of inequity aversion. Convergences of allocations and excesses are established via stochastic stability theory.

Index Terms— mean-field games; consensus; multiagent systems; network flow.

1. INTRODUCTION

In this paper we consider infinite copies of a coalitional game with transferable utilities (TU game). For each game, a central planner allocates revenues in order to stabilize the grand coalition, which occurs when the total amount given to the member of any sub-coalition exceeds the value of the sub-coalition itself (see the notion of "core" in [10]). In a continuous-time repeated game, the excess of a coalition is the cumulative discrepancy between the total amount given to the coalition and the value of the coalition up to the current time. The coalition's values are unknown but bounded, thus the excesses evolve according to controlled uncertain stochastic differential equations. The objective of the planner is to align the excesses with the average value computed over infinite copies of the same game. Such a phenomenon is known as crowd-seeking behavior in mean-field games.

Main result. For the problem at hand, we provide a mean-field game formulation and conduct a heuristic robust control design based on augmentation and regularization of the state space [6]. The mean-field game involves a macroscopic description based on a classical forward Kolmogorov partial differential equation which generates the distribution of the excesses over the horizon.

Related literature. The theory on mean-field games originated in the work of M.Y. Huang, P. E. Caines and R. Malhamé [7, 8] and independently in that of J. M. Lasry and P.L. Lions [9], where the now standard terminology of Mean Field Game (MFG) was introduced. The problem we analyze in this paper follows in spirit the study on robust dynamical TU coalitional games in [5] with the additional mean-field interactions between infinite copies of the same game, which was not present in [5]. Explicit solutions in terms of mean-field equilibria are not common unless the problem has a linear-quadratic structure, see [1]. This justifies our solution approach which approximates the original problem by an augmented linear quadratic one.

The rest of the paper is organized as follows. In Section 2, we illustrate the problem and introduce the model. In Section 3, we present the mean-field game. In Section 4, we illustrate the solution approach. Finally, in Section 5, we draw some conclusions and discuss future works.

Notation. Given a set $N = \{1, \dots, n\}$ of players and a function $\eta : S \mapsto \mathbb{R}$ defined for each nonempty coalition $S \subseteq N$, we write $< N, \eta >$ to denote the transferable utility (TU) game with players' set N and characteristic function η. We let η_S be the value $\eta(S)$ of the characteristic function η associated with a nonempty coalition $S \subseteq N$. Given a TU game $< N, \eta >$, we use $C(\eta)$ to denote the core of the game:

$$C(\eta) = \Big\{ x \in \mathbb{R}^n \,\Big|\, \textstyle\sum_{i \in N} x_i = \eta_N,\ \sum_{i \in S} x_i \geq \eta_S \text{ for all nonempty } S \subset N \Big\}.$$

Also, $\mathbb{R}_+$ denotes the set of nonnegative real numbers. Given a random vector ξ, the notation $\mathbb{E}[\xi]$ denotes its expected value. Given a Brownian motion (with drift) $\mathcal{B}(t)$, we denote

The work of D. Bauso was supported by the 2012 "Research Fellow" Program of the Dipartimento di Matematica, Università di Trento and by PRIN 20103S5RN3 "Robust decision making in markets and organizations, 2013-2016". The work of T. Başar was supported in part by the U.S. Air Force Office of Scientific Research (AFOSR) MURI grant FA9550-10-1-0573. D. Bauso is currently academic visitor at the Department of Engineering Science, University of Oxford, UK.

by $d\mathcal{B}(t)$ its infinitesimal increment, i.e., $\mathcal{B}(t) = \int_0^t d\mathcal{B}(\tau)$, the latter being the Itō integral. We use $\bar{\mathcal{B}}(t) = \frac{\mathcal{B}(t)}{t}$ to indicate the average infinitesimal up to time t. If $a(t)$ is the derivative of an almost everywhere differentiable function, the symbol $\tilde{a}(t)$ denotes the function itself, i.e., $\tilde{a}(t) = \int_0^t a(\tau)d\tau$. We also use $\bar{a}(t) = \frac{\tilde{a}(t)}{t}$ to indicate the average up to time t.

2. TU GAMES AS NETWORKS

Consider infinite copies of an n-palyer robust dynamical TU game $< N, \eta(t) >$, where the characteristic function $\eta(t)$ is a diffusion process with drift, whose evolution is described by the stochastic differential equation:

$$\begin{cases} d\eta(t) = w(t)dt + \sigma d\mathcal{B}(t), \;\; \text{in } \mathbb{R}^q, \\ \eta(0) = \eta_0, \end{cases} \tag{1}$$

where $q = 2^{n-1}$ is the number of coalitions. For each game, let a corresponding hypergraph $\mathcal{H}$ be given with vertex set V and edge set E as:

$$\mathcal{H} := \{V, E\}, \quad V = \{\mathbf{v}_1, \ldots, \mathbf{v}_q\}, \quad E := \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}.$$

The vertex set V has one vertex per coalition whereas the edge set E has one edge per player. A generic edge i is incident on

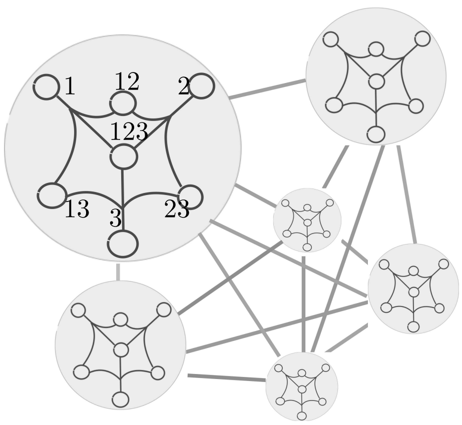

Fig. 1. Infinite copies of hypergraph $\mathcal{H} := \{V, E\}$ for a 3-player coalitional game.

a vertex $\mathbf{v}_j$ if the player i is in the coalition associated with $\mathbf{v}_j$. Thus, incidence relations can described by a matrix B whose rows are the characteristic vectors $c^S \in \mathbb{R}^n$. We recall that the components of a characteristic vector are $c_i^S = 1$ if $i \in S$ and $c_i^S = 0$ if $i \notin S$. Figure 1 depicts an example of hypergraph for a 3-player coalitional game on every single grey node. In the same spirit as in [5], we view allocation $u_i(t)$ as the flow on edge $\mathbf{e}_i$ and the coalition value $w_S(t)$ of a generic coalition S as the demand in the corresponding vertex $\mathbf{v}_j$. In view of this, allocation in the *core* of the game $C(\eta(t))$ translates into over-satisfying the demand at the vertices. Specifically,

$$\tilde{u}(t) \in C(\eta(t)) \quad \Leftrightarrow \quad B_{\mathcal{H}}\tilde{u}(t) \geq \eta(t), \tag{2}$$

with the last inequality satisfied as an equality due to the efficiency condition of the core, i.e, $\sum_{i=1}^n \tilde{u}_i(t) = \eta_q(t)$, where $\eta_q(t)$ denotes the qth component of $\eta(t)$ and is equal to the grand coalition value $\eta_N(t)$. Let $x(t) \in \mathbb{R}^q$ represent the coalition excess, whose time evolution is given by:

$$\begin{cases} dx(t) = (Bu(t) - w(t))dt - \sigma d\mathcal{B}(t), \\ x(0) = x_0. \end{cases} \tag{3}$$

We assume that the control and the disturbance are bounded within polytopes, i.e.,

- $u(t)$ is in the control set $U \subseteq \mathbb{R}^p$, $p > 0$,
- $w(t)$ is in the disturbance set $W \subseteq \mathbb{R}^q$, $q > 0$.

Given infinite copies of the same game, we can consider a probability density function $m : \mathbb{R}^q \times [0, +\infty[\to [0, +\infty[$, $(x, t) \mapsto m(x, t)$, which satisfies $\int_{\mathbb{R}^q} m(x, t)dx = 1$ for every time t. Let us also define the mean distribution at time t as $\bar{m}(t) := \int_{\mathbb{R}^q} x m(x, t)dx$.

In each game, the designer follows the so-called crowd-seeking law in that it adjusts the current allocation based on the average distribution of the other games.

Then, for each game, consider a running cost $g : \mathbb{R}^q \times \mathbb{R}^q \to [0, +\infty[$, $(x, \bar{m}) \mapsto g(x, \bar{m})$ of the quadratic form:

$$g(x, \bar{m}) \;\; = \;\; \frac{1}{2}\Big[(\bar{m} - x)^T Q \, (\bar{m} - x) \Big], \tag{4}$$

where $Q > 0$, that is positive definite.

Also consider a terminal cost $\Psi : \mathbb{R}^q \times \mathbb{R}^q \to [0, +\infty[$, $(x, \bar{m}) \mapsto \Psi(x, \bar{m})$ of the form

$$\Psi(x, \bar{m}) = \frac{1}{2}(\bar{m} - x)^T S(\bar{m} - x), \tag{5}$$

where $S > 0$. The problem in its generic form is then the following:

Problem 1 *Find the closed-loop optimal control and worst-case disturbance for the problem:*

$$\begin{cases} \inf_{u(\cdot)\in\mathcal{U}} \sup_{w(\cdot)\in\mathcal{W}} \Big\{ J(x_0, u(\cdot), w(\cdot), m(\cdot)) \\ = \mathbb{E}\Big[\int_0^T g(x(t), \bar{m}(t))dt + \Psi(x(T), \bar{m}(T)) \Big] \Big\}, \\ dx(t) = (B_{\mathcal{H}}u(t) - w(t))dt - \sigma d\mathcal{B}(t), \end{cases} \tag{6}$$

where $\mathcal{U}$ and $\mathcal{W}$ are the sets of all measurable functions $u(\cdot)$ and $w(\cdot)$ from $[0, +\infty[$ to U and W, respectively, and $m(\cdot)$ as a time-dependent function is the evolution of the distribution under the optimal control and the worst-case disturbance.

3. THE MEAN FIELD GAME

Let us denote by $v(x, t)$ the (upper) value of the robust optimization problem under worst-case disturbance starting from time t at state x (which in this case also turns out to be the lower value, and hence the *value*, since Isaacs condition [2] holds–see below). Problem 1 results in the following mean-field game system for the unknown functions $v(x,t)$, and $m(x,t)$:

$$\begin{cases} \partial_t v(x,t) + \inf_{u \in U} \sup_{w \in W} \Big\{ (Bu - w)^T \partial_x v(x,t) \\ +g(x,\bar{m},u,w) \Big\} + \frac{\sigma^2}{2} Tr\Big(\partial^2_{xx} v(x,t)\Big) = 0 \\ \qquad\qquad\qquad\qquad \text{in } \mathbb{R}^q \times [0,T[, \\ v(x,T) = \Psi(x,\bar{m}) \; \forall \, x \in \mathbb{R}^q, \\ \partial_t m(x,t) + div(m(x,t) \cdot (Bu - w)) \\ \qquad -\frac{\sigma^2}{2} Tr(\partial^2_{xx} m(x,t)) = 0, \text{ in } \mathbb{R}^q \times [0,T[, \\ m(0) = m_0, \\ \frac{d}{dt}\bar{m}_t = B\bar{u}^*_t - \bar{w}^*_t, \text{ in } [0,T[, \end{cases} \tag{7}$$

where $u^*(t,x)$ and $w^*(t,x)$ are the optimal time-varying state-feedback controls and disturbances, respectively, obtained as

$$\begin{cases} u^*(t,x) \in \arg\min_{u \in U}\{(Bu - w^*)\partial_x v(x,t) \\ \qquad\qquad +g(x,\bar{m})\}, \\ w^*(t,x) \in \arg\max_{w \in W}\{(Bu - w)\partial_x v(x,t) \\ \qquad\qquad +g(x,\bar{m})\}. \end{cases} \tag{8}$$

Note that the minimization and maximization problems above are completely decoupled, and hence in (7) the $\inf \sup$ is the same as $\sup \inf$ (that is, Isaacs condition holds [2]), and furthermore $\inf$ and $\sup$ can be replaced by $\min$ and $\max$, respectively, because of optimization of linear functions over closed and bounded finite-dimensional sets.

The first equation in (7) is the HJBI equation with variable $v(x,t)$. Given the boundary condition on final state (second equation in (7)), and assuming a given population behavior captured by $m(\cdot)$, the HJBI equation is solved backwards and returns the value function and best-response behavior of the individuals (first equation in (8)) as well as the worst adversarial response (second equation in (8)). The HJBI equation is coupled with a second PDE, known as the *Fokker-Planck-Kolmogorov (FPK)* equation (third equation in (7)), defined on variable $m(\cdot)$. Given the boundary condition on initial distribution $m(0) = m_0$ (fourth equation in (7)), and assuming a given individual behavior described by u^*, the FPK equation is solved forward and returns the population behavior time evolution $m(t)$. The last equation in (7) is obtained by averaging the left and right hand side of the dynamics (3). Any solution of the above system of equations along with (8) is referred to as *worst-disturbance feedback mean-field equilibrium*.

4. AUGMENTATION AND REGULARIZATION

This section illustrates a simple heuristic approach toward solving the set of equations (7)), based on state space augmentation and regularization [6]. The augmented state space includes the mean distribution, thus the augmented state variables evolve according to the equations

$$\begin{bmatrix} dx(t) \\ d\bar{m}(t) \end{bmatrix} = \Big(B \begin{bmatrix} u^*(x,t) \\ \bar{u}^*(t) \end{bmatrix} - \begin{bmatrix} w^*(x,t) \\ \bar{w}^*(t) \end{bmatrix} \Big) dt + \begin{bmatrix} \sigma d\mathcal{B}_t \\ 0 \end{bmatrix}. \tag{9}$$

For this system we introduce an assumption on the rate of convergence of the state $\bar{m}(t)$.

Assumption 1 *There exists a scalar $\theta > 0$ such that*

$$\frac{d}{dt}\bar{m}(t) = B\bar{u}^*(t) - \bar{w}^*(t) \geq -\theta \bar{m}_t, \textit{for all } t \in [0,T],$$

where the inequality is to be interpreted component-wise.

The above assumption implies that there exists a variable $\tilde{m}(t)$ which approximates the average mean value from below and evolves according to

$$\begin{cases} \frac{d}{dt}\tilde{m}(t) = -\theta \tilde{m}(t), \quad \text{for all } t \in [0,T], \\ \tilde{m}_0 = \bar{m}_0. \end{cases} \tag{10}$$

By substituting the current mean value $\bar{m}_t$ by its estimate $\tilde{m}_t$ the augmented problem is

$$\inf_{u(\cdot)\in\mathcal{U}} \sup_{w(\cdot)\in\mathcal{W}} \int_0^T \frac{1}{2}\Big[(\tilde{m}(t) - x(t))^T Q\, (\tilde{m}(t) - x(t)) \Big] dt$$

$$\begin{bmatrix} dx(t) \\ d\tilde{m}(t) \end{bmatrix} = \Big(\begin{bmatrix} 0 & 0 \\ 0 & -\theta I \end{bmatrix} \begin{bmatrix} x(t) \\ \tilde{m}(t) \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u(t) - \begin{bmatrix} I \\ 0 \end{bmatrix} w(t) \Big) dt + \begin{bmatrix} \sigma d\mathcal{B}(t) \\ 0 \end{bmatrix}.$$

Reformulating the problem in terms of the augmented state

$$X(t) = \begin{bmatrix} x(t) \\ \tilde{m}(t) \end{bmatrix},$$

and regularizing the solution via quadratic penalty terms on control and disturbance, we have the linear quadratic problem:

$$\inf_{u(\cdot)\in\mathcal{U}} \sup_{w(\cdot)\in\mathcal{W}} \int_0^T \Big[\frac{1}{2}(X(t)^T \tilde{Q} X(t) + u^T(t) R u(t) - w^T(t)\Gamma w(t))\Big] dt + \tilde{\Psi}(X(T))$$

$$dX(t) = \Big(FX(t) + Gu(t) + Hw(t) \Big) dt + L d\mathcal{B}_t,$$

where

$$\tilde{Q} = \begin{bmatrix} Q & -Q \\ -Q & Q \end{bmatrix}, \quad L = \begin{bmatrix} \sigma I \\ 0 \end{bmatrix},$$

$$F = \begin{bmatrix} 0 & 0 \\ 0 & -\theta I \end{bmatrix}, \quad G = \begin{bmatrix} B \\ 0 \end{bmatrix}, \quad H = \begin{bmatrix} -I \\ 0 \end{bmatrix},$$

$R > 0, \Gamma > 0$, and $\tilde{\Psi}(X) := \Psi(x, \tilde{m})$.
The idea is therefore to consider a new value function $\mathcal{V}_t(x, \tilde{m})$ (in compact form $\mathcal{V}_t(X)$) in the augmented state space which satisfies

$$\left\{ \begin{array}{l} \partial_t \mathcal{V}_t(X) + H(X, \partial_X \mathcal{V}_t(X)) \\ +\frac{1}{2}\sigma^2 \mathrm{Tr} \partial^2_{xx} \mathcal{V}_t(X) = 0, \text{ in } \mathbb{R}^{2q} \times [0, T[, \\ \mathcal{V}_T(X) = \tilde{\Psi}(X) \quad \text{in } \mathbb{R}^{2q}, \end{array} \right.$$

where $H(X, \partial_X \mathcal{V}_t(X))$ is the robust Hamiltonian [4]:

$$\begin{array}{l} H(X, \partial_X \mathcal{V}_t(X)) = \frac{1}{2} X^T \tilde{Q} X + \partial_X \mathcal{V}_t(X) F X \\ \quad -\frac{1}{2} \partial_X \mathcal{V}_t(X)[G R^{-1} G^T - H \Gamma^{-1} H^T](\partial_X \mathcal{V}_t(X))^T . \end{array}$$

This PDE admits the unique solution given by

$$\mathcal{V}_t(X) = \frac{1}{2} X(t)^T \underbrace{\left[\begin{array}{cc} P_{11}(t) & P_{12}(t) \\ P_{12}(t)^T & P_{22}(t) \end{array} \right]}_{P(t)} X(t) + \frac{1}{2} p(t),$$

where the symmetric matrix $P(t)$ satisfies (is the unique nonnegative-definite solution of) the generalized (game) Riccati differential equation

$$\begin{array}{l} \dot{P}(t) + P(t)F + F^T P(t) \\ -P(t)(G R^{-1} G^T - H \Gamma^{-1} H^T) P(t) + \tilde{Q} = 0 \,, \\ P(T) = \left[\begin{array}{cc} S & -S \\ -S & S \end{array} \right], \end{array} \quad (11)$$

and $p(\cdot)$ is solved from

$$\dot{p}(t) + \sigma^2 \, \mathrm{Tr} P(t) \,, \quad p(T) = 0 \,.$$

Given upper and lower bounds, u_i^+ and u_i^- respectively, let us introduce the sat function as in [3]:

$$sat_{[u_i^-, u_i^+]}\{\xi\} \doteq \left\{ \begin{array}{lll} u_i^- & \text{if} & \xi < u_i^- \\ u_i^+ & \text{if} & \xi > u_i^+ \\ \xi & \text{if} & u_i^- \leq \xi \leq u_i^+ \end{array} \right. .$$

Then a sub-optimal control is given by

$$\begin{array}{l} \tilde{u}(t) = sat\Big\{ - R^{-1} G^T P(t) X(t) \Big\} \\ = sat\Big\{ - R^{-1} B^T (P_{11}(t) x(t) + P_{12}(t) \bar{m}(t)) \Big\}, \end{array}$$

and the worst-case disturbance can be approximated by

$$\begin{array}{l} \tilde{w}(t) = sat\Big\{ \Gamma^{-1} H^T P X(t) \Big\} \\ = sat\Big\{ - \Gamma^{-1} (P_{11}(t) x(t) + P_{12}(t) \bar{m}(t)) \Big\}. \end{array}$$

The underlying idea of the approximation above is to consider the solution of the soft-constrained linear quadratic problem when the hard constraints are not active, while saturate every single component as soon as it reaches its upper or lower bound.

5. CONCLUSIONS AND FUTURE DIRECTIONS

We have provided a mean-field game formulation of infinite copies of "small worlds" each one described as a TU coalitional game. The problem has connections to recent research on robust dynamic coalitional TU games [5] and robust mean-field games [4, 6]. A quantitative analysis of the approximation error of the solution presented is left as future work.

6. REFERENCES

[1] M. Bardi, "Explicit solutions of some linear-quadratic mean field games," *Network and Heterogeneous Media*, 7, pp. 243–261 (2012).

[2] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory*, SIAM Series in Classics in Applied Mathematics, Philadelphia (1999).

[3] D. Bauso, F. Blanchini, and R. Pesenti, "Robust control strategies for multi inventory systems with average flow constraints," *Automatica*, vol. 42, pp. 1255–1266 (2006).

[4] D. Bauso, H. Tembine, and T. Başar, "Robust mean field games with application to production of an exhaustible resource," *Proceedings of 7th IFAC Symposium on Robust Control Design*, Aalborg, Denmark (2012).

[5] D. Bauso and P. Viswanadha Reddy and T. Başar, "Lyapunov stochastic stability and control of robust dynamic coalitional games with transferable utilities," submitted and under review (2012).

[6] D. Bauso and T. Mylvaganam and A. Astolfi, "Approximate solutions for crowd-averse robust mean-field games," submitted and under review (2013).

[7] M.Y. Huang, P.E. Caines, and R.P. Malhamé, "Individual and mass behaviour in large population stochastic wireless power control problems: centralized and Nash equilibrium solutions," *Proc. IEEE Conf. on Decision and Control*, HI, USA, December, pp. 98–103 (2003).

[8] M.Y. Huang, P.E. Caines, and R.P. Malhamé, "Large population cost-coupled LQG problems with non-uniform agents: individual-mass behaviour and decentralized ϵ-Nash equilibria," *IEEE Trans. on Automatic Control*, vol. 52. no. 9, pp. 1560–1571 (2007).

[9] J.-M. Lasry and P.-L. Lions, "Mean field games," *Jpn. J. Math.*, vol. 2, pp. 229–260 (2007).

[10] J. von Neumann, O. Morgenstern, Theory of Games and Economic Behavior, Princeton Univ. Press. (1944).