Dynamic Resource Allocation and Balanced Cell Loading - a Stochastic Meanfield Control Approach

Abheek Saha
Hughes Systique Corporation,
India
Email: abheek.saha@hsc.com

***Abstract*—Mean-field theory is a significant recent step for the field of stochastic optimal control. By allowing the optimal control functions to take into account not only the state of the controlled node, but also the mean-field state of an entire ensemble of nodes, mean-field theory allows us to model inter-dependent networks of agents in an analytically tractable manner. In this paper, we show its application to a very standard problem of cellular network optimization, the cell loading problem. By modelling the cell-loading problem as a combination of the loading of the individual cell, as well as the loading of the entire network, we show that a distributed optimal control function exists that can be individually implemented at nodes, and that is capable of reaching network wide equilibrium.**

***Keywords*—mean field games; stochastic control; cell loading; admission control**

I. Introduction and problem statement

Stochastic optimal control is a powerful technique to control time-varying systems with randomly varying inputs. Developed over the last fifty years from the base of variational inequality and deterministic optimal control theory, it has been applied in multiple disciplines, ranging from finance to oil exploration and medical trials. The fundamental strength of optimal control is the ability to develop an optimal control function which can optimize performance over a time interval, as opposed to a single instant of time, in the face of unknown, time varying inputs.

Application of stochastic optimal control to wireless networks, however, has been limited [1][2][3]. A fundamental problem in the application of optimal control techniques in this domain is that of inter-node inter-dependency. Wireless networks of the 4th and 5th generation are increasingly built around the principles of shared resources, overlapping coverage areas and inter-cell and even inter-radio technology coordination. This change, from the days of isolated cells of fixed boundaries in 2nd generation networks, has come about because of two reasons. The first is the ability of individual user terminals to use larger and larger bands of spectrum. The second is the need for networks to dynamically adapt to large variations in demand, both spatially and temporally. Cellular networks are being moved towards newer and newer business cases such as wide-area connectivity for cellular networks supporting *Internet Of Things*, connected vehicles, etc. Most of these use cases are dependent on network nodes being able to flexibly adapt to new patterns in user behaviour. Hence, the paradigm of dynamically shared resources and network node cooperation is here to stay. For a few examples, we see Coordinated Multipoint networks in 4G, Hetnets and Inter-Cell Interference Coordination (ICIC/eICIC). Indeed, the 3rd Generation Partnership Project (3gPP) has introduced the X-interface between network nodes as an explicit means of inter-node coordination in real-time, in order to make coordinated cooperative network operation possible.

A. Optimal Control for Wireless Networks

Applying any kind of optimal control to wireless network nodes hence needs to model the network nodes impact on each other. Network nodes are independent, yet coexisting agents, tied together by the constraints of shared resources and shared environments. In this situation, it is not really possible to model each network as independent of the other nodes. To apply optimal control, one would to simultaneously solve the optimal control equation for all network nodes simultaneously, i.e., the network state becomes a vector of states, one for each agent. This however leads to the dimensionality problem as the number of degrees of freedom increase as $\mathcal{O}(n^2)$. It also requires a degree of simultaneous coordinated control that is not feasible in most wireless networks. A strictly adversarial approach (such as used in game theoretic techniques) is also not appropriate, since network nodes are not necessarily operating in competition of each other. It may make sense for a given node to hand over load to another node or to take over loading from another node cooperatively. To a large extent, we are optimizing overall network capacity, not individual node capacity.

B. The Mean-Field Extension to Stochastic Optimal Control

In the 2000s, Lasry and Lions [4] and independently, yet nearly simultaneously Minyi Huang and his team [5] kicked off a concerted research effort on optimal control of multiple interacting stochastic processes with mean field constraints. Optimal control problems of this nature are called Mean Field Games (henceforth MFG). The MFG technique is an extension to stochastic optimal control that allows the empirical distribution of individual network node states to be included in the transition and cost functions. This provides us a mechanism for incorporating the network state variables into individual node decision control algorithms. For example, Huang et al. in [6] use mean-field stochastic control as a way of optimal power control in wireless networks. Wireless nodes have to set transmission power so as to maximize

 ISBN: 978-1-61208-650-7

the Signal to Interference Ratio (SIR), yet minimize cross-neighbour interference. In this case, the latter is modelled in terms of the empirical power distribution across the network.

In this paper, we apply stochastic control with mean field constraints to an associated problem, that of cell loading. We will show how this powerful new technique can be applied to this crucial and very basic problem of cellular resource management. The rest of the paper is organized as follows. In Section II, we introduce the cell loading problem. In Section III, we give an introduction to Stochastic Optimal Control and its extension to Mean-field constraints. Finally, in Section IV, we show how we model the cell loading problem in terms of mean-field constraints and stochastic demand and provide a framework for its solution.

II. The cell loading problem

The cell-loading problem has been studied as part of the load balancing problem since a long time and is seen as a fundamental component of the Self Optimizing Network (SON) [7][8].

A relatively recent analysis of the current status and open areas is given by Andrews et al. in [9]. In this work, the authors also discuss the myths surrounding cell loading and QoS. One of the myths identified by the authors is that the capacity of a cell is rarely a property of the link SIR, but also has to take into account the loading of the cell itself. In our opinion, this underlies the need to do active load balancing as discussed in the rest of the paper.

1) Problem Description: The problem is briefly described as follows: we have a network of overlapping cells covering a given coverage area. Each cell is controlled by a network node (base-station). The state variable $X(t)$ for a given cell is the demand for bandwidth from the associated network node. The network nodes negotiate with a central controller for allocation of resources; the resources available to a given network are a measure of its capacity $c(t)$. The resources allocated to a cell (network node) may be a combination of various different physical and computational resources, such as spectrum, power and backhaul capacity. All of these combine in some way to determine the overall load handling ability $c()$ of a given network node. Since these are shared resources, nodes in a SON can flexibly deploy them, while keeping with overall network constraints, from one cell to the other as demand changes.

We make the problem more interesting by making some additional assumptions on the part of the users; that they are bandwidth hunting and self-optimizing. A bandwidth hunting entity constantly increases demand as its current demand is met; this attribute is typically used for TCP congestion management algorithms, which hunt for spare bandwidth in the network and then fill it up. Since modern wireless networks are dominated by data traffic, this is not a far-fetched assumption. The second attribute refers to the user equipments agency in terms of selecting the cell to attach to; an individual user terminal will tend to detach from over-crowded cells with less available bandwidth and attach to less crowded cells using a mixture of measurements and network feedback. 5th generation user terminals will have the capability to interrogate the network for this kind of information and the algorithms to use the information for optimal network selection.

It is clear that the likelihood of input demand rising further is tied to the expressed demand not only in the current cell, but also in the neighbouring cells; for example, if the demand for a given cell is low and that of the neighbouring cells higher, it is possible for users to handoff to one of the neighbouring cells, hence decreasing expressed demand in the given cell. Here we can use the difference of the local demand and the average of the empirical network-wide demand to express this preference. Similarly, we can place constraints on the final distribution, by making $g()$ a function both of the final value of X, as well as the final distribution; for example, by providing an incentive for users to stay within a certain range of the mean.

Given this scenario, the challenge is to design an algorithm for the optimal allocation of resources. The input to the algorithm is the demand as measured at each network node, and information of the network wide distribution of this variable; since we are only interested in a single moment of the distribution, the amount of information to be circulated network wide is relatively limited. Based on this input, individual network nodes will compute the optimal capacity they wish to deploy and then execute that strategy. The objective is to minimize the demand allocation gap while maximizing the total served demand. For reasons we shall describe below, we shall formulate and solve the algorithm as a stochastic optimal control problem with a meanfield constraint.

2) Previous Work: The existing literature in load balancing in cellular networks is vast, even if we limit it to distributed cooperative algorithms. Broadly, the approaches in the literature can be divided into two categories. One set of research tends to focus on user redistribution, using intra-cellular and inter-cellular handoffs [10][11]. In other words, rather than moving resources, the users are moved between cells. In these approaches, the decisions are typically taken at the endpoints with the network nodes providing information about current loading. Alternately, one can move the decision logic to the network nodes themselves. In the second set of approaches, the network nodes autonomously learn the *optimal* loading limit individually and then act to achieve this. In [12], the authors propose reinforcement learning techniques for network nodes to tune specific configuration parameters to achieve the optimal load. In contrast, in [13] Bigham et al. describe a method of structured direct negotiation between the network nodes, using the gap between demand and capacity as a distance factor between nodes in a graph. In [14], the authors model the negotiation process as a game between an individual loaded cell and underloaded neighbour cells.

In this paper we have proposed load balancing using mean-field stochastic optimal control. We replace the inter-network node negotiation by a distributed stochastic optimization process, where the network node has visibility of its own load as well as the network wide load. The network wide load is a *mean field* statistic, sampled at a central location from feedback from individual nodes, computed, filtered and broadcast back to the network nodes as an input variable. User load is modelled as an exogenous variable, where individual users seek to maximize their own utilities.

III. Stochastic Optimal Control and Mean Field Games

In this section, we shall present the basics of mean-field stochastic optimal control.

A. Fundamentals

We start with the basic stochastic optimal control problem. We consider a system whose state variable X is controlled by the transition function (1) given below

$$dX_t^k = b(X_t^k, u_t)dt + \sigma(X_t^k, u_t)dW_t^k \quad (1)$$

The variable $u_t = U(X_t^k, t)$ is the output of a control function at time t, where said function is adapted to the filtration generated by the stochastic process W_t^k , which is a brownian motion. $b()$ and $\sigma()$ are Lipschitz continuous bounded functions as required for the standard definition of a Wiener process. The system governed by this equation has a long term cost function as in (2).

$$\begin{aligned}\Phi\left(x^0, u, T\right) &= \mathbb{E}\left[g(X_T) + \int_0^T f\left(X_t^k, u_t\right) dt\right] \\ X(0) &= x^0\end{aligned} \quad (2)$$

Our aim is to find the optimal control function $u^*(t)$ from a set of possible control functions $u \in \mathscr{U}$ so as to minimize the expected minimum total cost $\Phi(x0, a, T)$, over the time period $[0, T]$. Computing the optimal value of $u(t)$ is, in a nutshell, the stochastic optimal control problem.

The general solution technique is derived from the corresponding deterministic optimal control problem, with an important caveat. Whereas, in a deterministic control problem, the cost corresponding to each choice of $u()$ can be forecast, here we are faced with uncertainty in the future. At each point t, the value of $u(t)$ has to be based on the information regarding X as known upto then. This is what we mean by being adapted to the filtration of X.

1) Solution Technique : HJB equation : The classic way to solve a stochastic optimal control problem is to construct the Hamilton Jacobi Bellman (HJB) equation, which, for the above problem is given in (3).

$$\begin{aligned}&\frac{\partial\phi}{\partial s}(y,u) + b(y,u)\nabla_x\phi + \frac{\sigma^2(y,u)}{2}\nabla_x^2\phi + f(y,u) = 0 \\ &\Rightarrow \frac{\partial\phi}{\partial s}(y,u) + \mathcal{H}(b, \nabla_x\phi, f, u) + \frac{\sigma^2}{2}(y,u)\nabla_x^2\phi = 0 \\ &\phi(Y) = g(Y)\end{aligned} \quad (3)$$

The value of $u = u^*$ which solves this equation for all y gives the optimal value of u. Note the second derivative term, which makes the solution rather complex. The function $\mathcal{H}(y, b, f, x, u) = \langle y(x, u), b(x, u)\rangle + f(x, u)$ is called the Hamiltonian; the solution to the above equation depends, to a very large extent, on the structure of the Hamiltonian.

2) The Adjoint Equation Approach: A second method is to construct the adjoint equation based on the Stochastic Maximum Principle (SMP), in a manner analogous to the Lagrangian for a deterministic optimization problem. The stochastic maximum principle [15] is conceptually similar to the Pontryagin maximum principle for the deterministic case. However, it is more complex to solve because the solution is not time reversible [16].

The SMP requires us to find two stochastic variables p_t, q_t, such that the equation pair (4) holds.

$$\begin{aligned}-dp_t &= \nabla_x\mathcal{H}\left(p_t, q_t, b, f, X_t^k, u\right)dt + q_t dW_t^k \\ p_T &= \nabla_x g\left(X_T\right)\end{aligned} \quad (4)$$

In this equation, the Hamiltonian takes the extended form defined as in (5).

$$\begin{aligned}&\mathcal{H}(p, q, b, f, x, u) \\ &\quad = \langle p, b(x, u)\rangle + \text{tr}\left\{q^T.\sigma(x, u)\right\} + f(x, u) \\ &\partial_u\mathcal{H}(p, q, b, f, x, u^*) = 0\end{aligned} \quad (5)$$

The solution of p_t, q_t if they exist, provide an optimal control function u. Note that this is a backward stochastic differential equation again because the termination value of p is provided. We note that this is a simplified version of the SMP, where the function $\sigma()$ is independent of X_t^k. If $\sigma()$ is a function of X, then we need to add a second pair of variables to take care of the additional risk of modifying the diffusion term in dX_t^k. The interested reader should consult Yong [16, Section 3.1] for more information. For the rest of this article, we will only consider problems where $\sigma()$ is independent of u.

B. Adding the Meanfield Constraint

We now consider the problem of adding the meanfield constraint. In this version of the problem, the equations (1), (2) change to the form given in (6). The term z_t is the meanfield term. In the simplest case, it is the scaled average of the empirical states of the other agents in the game, i.e., $z_t = \eta/(N-1)\sum X_t^{j\neq k}$. In real life, it can be expressed as more complex moments of the empirical distribution μ_t^X or functions thereof.

$$\begin{aligned}dX_t^k &= b(X_t^k, u_t, z_t^k)dt + \sigma(X_t^k, u_t, z_t)dW_t^k \\ \phi\left(x^0, u\right) &= E\{g\left(X_T, z_t\right) + \int_0^T f\left(X_t^k, u_t, z_t\right)dt\} \\ \mu_t^X(Y) &= \frac{1}{N}\sum_{j=1}^N \mathcal{I}_{X_j=Y} \\ z_t &= \text{Average}\left(\mu_t^X\right)\end{aligned} \quad (6)$$

The existence of the term z_t represents the coupling between the states of the different agents. It requires us to take into account the global ensemble of states, when computing the optimal strategy $u^*()$ for the kth agent. The meanfield term thus allows us to model the interdependence of the agents. Most explicit solutions that we have encountered use the empirical average. However, more complex functions may also be used at the cost of complexity.

1) Convergence to Equilibrium: The incorporation of the meanfield term μ_t^X raises an interesting problem of evolution of the meanfield distribution μ_t^X in response to a given strategy $u()$. This is important because we would like a solution where the optimal strategy $u^*()$ drives μ_t^X to stable equilibrium (or at least a stable value of the feedback term z_t). Huang et al. in [5] address this problem by considering the case where the number of agents is very large. By taking the limit to infinity, Huang et al. demonstrate that there is a Nash equivalent solution (NCE) wherebe μ_t^X tends to a long term stable distribution μ_t^X which leads to a Nash equilibrium for all agents. This is very important, because it lets us relate the term μ_t^X to the evolution of X; we shall see the solution technique in The HJB-KFP approach section below.

A second interesting problem is that of differentiating the Hamiltonian function with respect to a distribution function; Lasry [4] has shown that this can be done using the Wasserstein space of probablity measures on a Borel space and using a suitably defined lifting function.

Unfortunately, solutions for stochastic optimal control problems with mean-field games are not easy. There are three main techniques, two of which depend on solving Forward Backward Stochastic Differential Equations (FBSDE). To date, most of the research in solutions of MFGs pertain to a special class of MFGs, the so-called Linear Quadratic MFG [17][18][19]. There are two main approaches that we shall discuss below; these approaches have been studied mostly in the context of LQMFGs. Recently, a paper has been published by Pham and Wei [20], which discusses a dynamic programming solution to these games. However, we have not covered it here.

The linear quadratic MFG (LQMFG) consists of an optimization problem where the transition function of X_t^k is linear (7) and the value function is quadratic (8).

$$dX_t = bX_t + au_t + \hat{b}z_t \tag{7}$$

$$\phi = qX^2 + r.u^2 + \hat{q}\left(X - z\right)^2 \tag{8}$$

While simple in nature, the LQMFG can be applied to a large number of situations with interesting results. As we shall see below, we have modelled the cell loading problem as an LQMFG.

2) The HJB-KFP approach: The fundamental idea behind this approach is that as the number of agents becomes large, the distribution for the states of the individual agents approaches the probability distribution for the state of each individual agents. In [5] Huang et al. have shown that this assumption leads to a Nash equilibrium. The solution comes from utilizing the Kolmogorov Backward equation (sometimes called the Kolmogorov Fokker Planck equation) to model the probability distribution of X for a given agents, together with the HJB equation, as shown below (9), given the probability distribution of the starting state. In theory, in a stable equilibrium, the long term probability distribution of X_t^k under the Fokker Plank equilibrium should match the empirical distribution of X_T^k as $T \to \infty$, leading to a stable solution for the HJB equation and thereby making the equilibrium self-sustaining. In this situation, we can postulate that $X_T^k \to z_t$ as the distribution evolves, for large values of T.

$$\begin{aligned}
&\partial_t \phi + \frac{\sigma^2}{2}\nabla^2 \phi + \mathcal{H}(\nabla_x \phi, b, f, X_t^k, u_t, z_t) \\
&\phi_T = g\left(X_T\right) \\
&z_t = \mathbb{E}\left[X_t^k\right],\ X_0^k = x_0 \\
&\partial_t z_t = -b()\nabla_x z_t + \frac{1}{2}\sigma^2(x)\nabla^2 z_t,\ z_0 = X_0^k
\end{aligned} \tag{9}$$

Note that the HJB equation is a backward stochastic differential equation, whereas the KFP is a forward equation. Once again, the value of u which solves both equations simultaneously is the optimal control function. The KFP-HJB technique has been used successfully for LQMFGs in many papers; a good example is that of Bardi [17].

C. *Constructing the Adjoint Equation*

An alternate approach to solve the mean field problem is to extend the adjoint equation described in (4) to take into account the presence of the mean field term z_t [21]. To do this, we have to extend the Hamiltonian as shown in (10) .

$$\tilde{\mathcal{H}}\left(X, y, z, \tilde{X}, u\right) = \mathcal{H}\left(X, y, z, \mu_t^X, u\right) \tag{10}$$

Here $\tilde{X}$ is a random variable with a probability distribution function matching μ_t^X. The extended Hamiltonian $\tilde{\mathcal{H}}$ thus becomes a *lifted* version of the standard Hamiltonian, allowing us to take the derivative with respect to the distribution μ_t^X. The Stochastic Maximum Principle is as in (11)

$$\begin{aligned}
-dp_t &= \nabla_x \mathcal{H}\left(X_t^k, u_t, p_t, q_t, \mu_t^X\right) dt \\
&\quad + \mathbb{E}\left[\partial_\mu \mathcal{H}(X_t^k, u_t, p_t, q_t, \mu_t^X)\right] + q_t dW_t^k \\
p_T &= \nabla_x g\left(X_T\right) + \mathbb{E}\left[\partial_\mu g\left(X_T\right)\right]
\end{aligned} \tag{11}$$

The challenge with solving (11) is that we have no idea of μ_t^X or even of the form of μ_t^X. One possible way out of this is to treat it as a variational inequality problem as suggested by Bensoussan in [18]. In this approach, we assume that all the agents, other than the kth agents is using the optimal strategy, which leads to the term μ_t^X being replaced by a deterministic z_t as the moment of the distribution that we are interested in. The optimal strategy, if it is deviation proof, will lead to a fixed point solution, where by $x_t = \mathbb{E}\left[X_t | g(z_t)\right] = z_t$, where $g(z_t)$ represents the moment of the distribution that we are interested in. Bensoussan et al. apply this to the solution of a Linear Quadratic MFG and demonstrate the solution is ϵ Nash compliant.

IV. Application to the cell-loading problem

We now see if this technique can be applied to the cell-loading problem. We recall that the purpose of the cell loading problem is to ensure that cell-loads are as uniform as possible, given the variations in demand. We measure the expressed demand ,i.e., actual request for service at any given kth cell as the state variable X_k. The variation in demand is based on two parts. One is the natural variation, captured through a random diffusion term. The second is the variation of demand in response to the offered service. For our control variable u, we have selected the gap between the request for service $X_k(t)$ and the actual capacity assigned to that cell, $C_k(t)$. $u_t^k = X_t^k - C_t^k$.

A. Feedback Loop Between Demand and Offered Capacity

The majority of modern data-based applications use the Internet Transport Control Protocol (TCP) as the backbone transport protocol. This is true for Internet browsing, as well as video streaming using Dynamic Adaptive Streaming over HTTP (DASH). TCP by its very design uses a bandwidth hunting algorithm to determine the appropriate transmission rate. As a result, TCP endpoints react to the available bandwidth in the network. When the network is congested, the TCP back off and reduce the data injection rate and hence, the network load. On the other hand, if they sense the availability of bandwidth in the network, they increase the network load gradually. The success of the TCP bandwidth hunting algorithm is such that even non TCP connections are nowadays required to maintain *TCP like* transmission rate management protocols. For example, the Tcp Friendly Rate Control [22] is now an Internet standard for bandwidth control of media flows such as those proposed in Web real-time communication (WebRTC). The assumption of bandwidth hunting endpoints is important, because it allows us to model demand as a continuous process. If, on the other hand, the bandwidth demand changed in discrete bands, we would have to use a jump-diffusion process, which makes the analysis more complex.

For wireless networks, the bandwidth hunting behaviour of individual endpoints is augmented by the bandwidth sensing capability of the access network user; for example, User Equipment (UE) triggered handoffs between cells as a response to congestion. We can postulate that expressed demand will decrease in the face of a demand capacity gap ($u > 0$) and increase in the face of surplus capacity being deployed ($u < 0$). We further postulate that this has to take into account the overall distribution of demand. In other words, if a given kth cell is heavily loaded and facing a demand supply gap, its users will have an incentive to migrate to neighbouring cells. Hence, we define the state transition function as in (12), using the term z_t as introduced above.

$$dX^k(t) = -Au_t^k + B\left(X_t^k - z_t\right) + \sigma dW_t^k \tag{12}$$

We note that the diffusion term is independent of the empirical distribution, for the reasons described above. A and B are constants.

B. Network Cost

We now come to the cost function. As expected, we penalize the admission control function for high values of u; if positive, because of the large demand supply gap and if negative, because of the oversupply and consequent wastage of capacity. The final reward function is purely a function of the empirical distribution of X.

$$f() = M.u_t^2 - NX_t^2 \tag{13}$$
$$g() = \eta\left(X_T - z_T\right)^2 \tag{14}$$

C. Existence of a Solution

We will now show that a unique solution exists, for suitable values of M, N, A and B. We can use the technique given in Bensoussan [18, Section 3], by taking the derivative of the cost function at the optimal u and then setting it to zero. The cost function can be written in terms of the perturbed optimal cost $u_t(\theta) = u_t^* + \theta v_t)$ as in (15).

$$\begin{aligned}
&\phi(u_t) = \eta\mathbb{E}\left[X_T - z_T\right]^2 + \mathbb{E}\left[\int_0^T \left(Mu_t^2 - NX_t^2\right)dt\right] \\
&dX_t^k = Au_t + B\left(X_t^k - z_t\right) \\
&X_t^k\left(u_t^* + \theta v_t\right) = y_t + \theta\tilde{x}_t,\ d\tilde{x}_t = Av_t + B\tilde{x}_t
\end{aligned} \tag{15}$$

Taking the derivative of $\phi(u_t + \theta v_t)$ with respect to θ and setting it to 0,

$$\mathbb{E}\left[\int_0^T 2Mu_tv_t - 2N\tilde{x}_t.y_tdt\right] + \mathbb{E}\left[\eta\tilde{x}_Ty_T\right] = 0 \tag{16}$$

We choose an adjoint variable ω_t with the properties shown in (17).

$$\frac{d\omega_t}{dt} = -B\omega_t + 2Ny_t,\ \omega_T = \eta y_T \tag{17}$$

Expanding $d(\omega_t\tilde{x}_t)$ and substituting suitable in (16), we get the following pair of adjoint equations.

$$\begin{aligned}
\frac{d\omega_t}{dt} &= -B\omega_t + 2Ny_t,\ \omega_T = \eta y_T \\
dy_t &= -\frac{A^2}{2M}\omega_t dt + B\left(y_t - z_t\right)dt + \sigma dW_t
\end{aligned} \tag{18}$$

The optimal u is given by (19). The form of $p(t)$ has to be chosen so that $u^*()$ is anticipatory (because ω_t is not guaranteed to be $\mathscr{F}(Y)$ adapted.

$$u_t^* = -A/2Mp(t), p(t) = \mathbb{E}\left[\omega_t|\mathscr{F}(Y)\right] \tag{19}$$

By the fixed point theorem, a solution exists if $z_t = [E]\left[y_t\right]$. Setting $\mathbb{E}\left[\omega_t\right] = \epsilon_t$ and $\mathbb{E}\left[y_t\right] = Y_t$, we can write the above equation as (20). It is clear that a fixed point solution exists, because the matrix is invertible.

$$\frac{d}{dt}\begin{bmatrix}\epsilon_t \\ Y_t\end{bmatrix} = -\begin{bmatrix}B & -2N \\ \frac{A^2}{2M} & 0\end{bmatrix}\begin{bmatrix}\epsilon_t \\ Y_t\end{bmatrix} \tag{20}$$

The equation pair in (18) has to be solved using numerical techniques [23]. In some rare cases an analytical solution is available. The output of the solution is an approximate control function $u(x, z)$, which is a function of the current state value x and the network state z. The control function may be stored as a two-dimensional table or as a polynomial function and computed at appropriate intervals. In a future paper, we shall present the challenges of solving the MFG and the associated performance for a large network of nodes.

V. Conclusion and Future Work

In this paper, we have demonstrated the application of mean-field stochastic optimal control to a very standard and well-studied problem of wireless network control. As we have seen here, even a simplified network model can capture a rich network interaction structure and yield a sophisticated, yet realizable solution to this problem. We have demonstrated that a solution exists for a simple linear format of the game, which can be solved numerically.

It is arguable that our particular model for the cell loading problem can be significantly enhanced. For example, we can put further constraints on the solution space. This may include domain specific constraints, i.e., a maximum limit on the capacity per cell or the total capacity in the network, etc. Since our primary purpose is to demonstrate the applicability of the Mean-field technique, we have used a simplified model in this paper for the sake of analytical tractability. Most existing solution techniques for MFGs are limited to very specific models. We hope to extend our work to more complete models as our ability to solve more complex MFGs evolves.

References

[1] L. Georgiadis, M. J. Neely, L. Tassiulas *et al.*, "Resource allocation and cross-layer control in wireless networks," Foundations and Trends in Networking, vol. 1, no. 1, 2006, pp. 1–144.

[2] M. J. Neely, E. Modiano, and C.-P. Li, "Fairness and Optimal Stochastic control for Heterogeneous Networks," IEEE/ACM Transactions On Networking, vol. 16, no. 2, 2008, pp. 396–409.

[3] T. Holliday, A. Goldsmith, N. Bambos, and P. Glynn, "Distributed power and admission control for time-varying wireless networks," in Information Theory, 2004. ISIT 2004. Proceedings. International Symposium on. IEEE, 2004, pp. 352–352.

[4] J.-M. Lasry and P.-L. Lions, "Mean field games," Japanese journal of mathematics, vol. 2, no. 1, 2007, pp. 229–260.

[5] M. Huang, R. P. Malhamé, P. E. Caines *et al.*, "Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle," Communications in Information & Systems, vol. 6, no. 3, 2006, pp. 221–252.

[6] M. 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," in Decision and Control, 2003. Proceedings. 42nd IEEE Conference on. IEEE, 2003, pp. 98–103.

[7] O. G. Aliu, A. Imran, M. A. Imran, and B. Evans, "A survey of self organisation in future cellular networks," IEEE Communications Surveys & Tutorials, vol. 15, no. 1, 2013, pp. 336–361.

[8] D. Liu, L. Wang, Y. Chen, M. Elkashlan, K.-K. Wong, R. Schober, and L. Hanzo, "User association in 5G networks: A survey and an outlook," IEEE Communications Surveys & Tutorials, vol. 18, no. 2, 2016, pp. 1018–1044.

[9] J. G. Andrews, S. Singh, Q. Ye, X. Lin, and H. S. Dhillon, "An overview of load balancing in hetnets: old myths and open problems," IEEE Wireless Communications, vol. 21, no. 2, April 2014, pp. 18–25.

[10] H. Kim, G. De Veciana, X. Yang, and M. Venkatachalam, "Distributed alpha-optimal user association and cell load balancing in wireless networks," IEEE/ACM Trans. Netw., vol. 20, no. 1, Feb. 2012, pp. 177–190. [Online]. Available: https://doi.org/10.1109/TNET.2011.2157937

[11] K. Son, S. Chong, and G. D. Veciana, "Dynamic association for load balancing and interference avoidance in multi-cell networks," IEEE Transactions on Wireless Communications, vol. 8, no. 7, July 2009, pp. 3566–3576.

[12] P. Muñoz, R. Barco, J. M. Ruiz-Avilés, I. de la Bandera, and A. Aguilar, "Fuzzy rule-based reinforcement learning for load balancing techniques in enterprise lte femtocells," IEEE Transactions on Vehicular Technology, vol. 62, no. 5, 2013, pp. 1962–1973.

[13] L. Du, J. Bigham, L. Cuthbert, P. Nahi, and C. Parini, "Intelligent cellular network load balancing using a cooperative negotiation approach," in 2003 IEEE Wireless Communications and Networking, 2003. WCNC 2003., March 2003, pp. 1675–1679 vol.3.

[14] A. Awada, B. Wegmann, I. Viering, and A. Klein, "A game-theoretic approach to load balancing in cellular radio networks," in Personal Indoor and Mobile Radio Communications (PIMRC), 2010 IEEE 21st International Symposium on. IEEE, 2010, pp. 1184–1189.

[15] S. Peng, "A general stochastic maximum principle for optimal control problems," SIAM Journal on control and optimization, vol. 28, no. 4, 1990, pp. 966–979.

[16] J. Yong and X. Y. Zhou, *Stochastic controls: Hamiltonian systems and HJB equations*. Springer Science & Business Media, 1999, vol. 43.

[17] M. Bardi, "Explicit solutions of some linear-quadratic Mean field games," Networks and heterogeneous media, vol. 7, no. 2, 2012, pp. 243–261.

[18] A. Bensoussan, K. Sung, S. C. P. Yam, and S.-P. Yung, "Linear-quadratic Mean field games," Journal of Optimization Theory and Applications, vol. 169, no. 2, 2016, pp. 496–529.

[19] O. Guéant, "Mean field games - equations with quadratic Hamiltonian: a specific approach," Mathematical Models and Methods in Applied Sciences, vol. 22, no. 09, 2012, p. 1250022.

[20] H. Pham and X. Wei, "Dynamic Programming for McKean-Vlasov dynamics," Control and Optimization, SIAM Journal on, vol. 55, no. 2, 2017, pp. 1069–1011.

[21] R. Carmona, F. Delarue *et al.*, "Forward–backward stochastic differential equations and controlled McKean–Vlasov dynamics," The Annals of Probability, vol. 43, no. 5, 2015, pp. 2647–2700.

[22] M. Handley, S. Floyd, J. Padhye, and J. Widmer, "TCP Friendly Rate Control (TFRC): Protocol Specification," Internet Requests for Comments, RFC Editor, RFC 3448, January 2003. [Online]. Available: http://www.rfc-editor.org/rfc/rfc3448.txt

[23] J. Douglas, J. Ma, P. Protter *et al.*, "Numerical methods for forward-backward stochastic differential equations," The Annals of Applied Probability, vol. 6, no. 3, 1996, pp. 940–968.