BOUNDEDLY-RATIONAL FAST-TUNING CONTROL THEORY AND STATISTICAL MECHANICS

MICHAEL CAMPBELL

Aurislink

E-mail address: michaeljcampbell@outlook.com.

Abstract. We construct a model of control theory with 'fast-tuning' of parameters relative to the ambient dynamics of the system. The parameters are tuned 'myopically' (i.e., small changes are made), along with a random perturbation that allows for a large net change with certain probability. This is modeled using a drift-diffusion stochastic partial differential equation. The idea is to model 'bounded rationality' of the agent(s) tuning the parameters – that is, they may not follow the optimal path for tuning because of a lack of complete information about the system, errors in judgement, and/or a desire to experiment and test other options.

Keywords: Control Theory, Bounded Rationality, Statistical Mechanics.

1. Introduction

One aspect of modeling is to use "real-world" constraints to attempt to simplify the model or at least to make it more tractable. For example, in studying a physical system with a very large numbers of particles such as a glass of water, it becomes intractable to analyze a number of equations of the order 10^{23}. The field of statistical mechanics was invented to make simplifying assumptions, but to still capture important aspects of the physical process being studied, such as changes of phase of the water. A common issue for control systems is complexity, and in many cases the agents [1] who tune the parameters do not have complete information about the system since it is too complex. As such, they make decisions that may be partially rational, but also have a non-rational component. Such non-rational components of decision making could be the use of 'rules-of-thumb', educated guesses, intuitive hunches, emotional preferences or biases, or even errors such as assumed 'rational' decisions that are made based on incomplete or incorrect information.

[1] Examples of agents are decision-making entities such people, machines, etc.

We consider the situation where agents have the ability to tune parameters very quickly relative to the ambient system dynamics. An example of this is the study of speculators in a two-market system with a long-term, large-scale investor [4, 5, 6]. In this model, speculators reach equilibrium during a duration of time for which the variables for the long-term investor dynamics are constant. Hence we have two timescales: a short one for the parameters, and a longer one for the ambient dynamics.

2. Control Theory Model

A classical control problem can be written as an optimization problem with an objective function to be optimized and with constraints on state variables (c.f. [14, 12]). We have state variables $\boldsymbol{x} = (x_1, \ldots, x_M) \in \mathbb{R}^M$ with each $x_i \in I_x \subset \mathbb{R}$, where I_x is an interval. The state domain is

$$\Omega_x = \prod_{i=0}^{M} I_x. \tag{1}$$

We also have parameters $\boldsymbol{u} = (u_1, \ldots, u_N) \in \mathbb{R}^N$ with each $u_i \in I_u \subset \mathbb{R}$, where I_u is a closed, bounded interval. The parameter domain is

$$\Omega_u = \prod_{i=0}^{N} I_u. \tag{2}$$

A **control** $\boldsymbol{\alpha} : [0, T] \to \Omega_u$ is a curve in parameter space Ω_u. For each control $\boldsymbol{\alpha}$, the state variables evolve according to the state equations which give the ambient dynamics:

$$\boldsymbol{x}(t) = \boldsymbol{x}(0) + \int_0^t \boldsymbol{f}(\tau, \boldsymbol{x}(\tau), \boldsymbol{\alpha}(\tau)) d\tau, \qquad \boldsymbol{x}(0) = \boldsymbol{x}_0, \;\; t \in [0, T]. \tag{3}$$

The curve $\boldsymbol{x}(t)$ is regarded as the response of the system. The main goal is to find a control that maximizes the objective function

$$\Phi(\boldsymbol{\alpha}) \;=\; \int_0^T g\left(\boldsymbol{x}(t), \boldsymbol{\alpha}(t)\right) dt + \Psi(\boldsymbol{x}(T)), \tag{4}$$

where g is the operational running cost function, and Ψ the operational terminal cost function, and $\boldsymbol{x}$ solves (3) for the control $\boldsymbol{\alpha}$.

Now we depart from classical control theory and implement two time scales. We keep the variable t for the slower ambient dynamics of the state variables $\boldsymbol{x}$, and the variable t' for the faster dynamics of the parameters $\boldsymbol{u}$.

For a given time t, the parameters are chosen in a boundedly-rational way to optimize the objective function at the specific time t. Because of the fast-tuning of agents, they only "see" the objective function at a specific time t and *maximize only the integrand* of (4). Hence we have a new **objective function**

$$\Phi_{\boldsymbol{x}(t)}(\boldsymbol{u}) \; = g(\boldsymbol{x}(t), \boldsymbol{u}) + \Psi(\boldsymbol{x}(T))\delta_T(t), \tag{5}$$

where t is constant (and thus $\boldsymbol{x}$ is also constant), and $\delta_T(t) = 0$ for $t \neq T$ and $\delta_T(T) = 1$. We point out that g above in (5) should have as argument a curve $\boldsymbol{u}(t')$ instead of just the variable $\boldsymbol{u}$. However, the curve $\boldsymbol{u}(t')$ that would maximize the objective function would be the constant curve $\boldsymbol{u}(t') = \boldsymbol{u}^*$, where $g(\boldsymbol{x}(t), \boldsymbol{u}^*) \geq g(\boldsymbol{x}(t), \boldsymbol{u})$ for all $\boldsymbol{u}$. Thus the control problem becomes one of maximizing the objective function J over variables $\boldsymbol{u}$, not curves $\boldsymbol{u}(t')$. This property is a result of the fast-tuning of parameters relative to the slow ambient dynamics.

Since agents are boundedly rational, we assume they optimize by following the gradient of the objective function, and this optimization is perturbed by a random error term. We also add a process that keeps the paths of the parameters within given bounds/constraints of the domain, with reflections off the boundary of the parameter domain (a "box") with respect to the normal vector at each face of the box (c.f. [15]).

The stochastic dynamics are given by the Itô diffusion (Langevin) equations: for $1 \leq i \leq N$,

$$du_i = \tfrac{\partial}{\partial u_i}\Phi_{\boldsymbol{x}}(\boldsymbol{u})dt' + \nu dw_i(t') + r_i(\boldsymbol{u})dq_i(t'), \tag{6}$$

where the $w_i(t')$ are zero-mean, unit-variance normal random variables from a Wiener process [2] ν is a variance parameter, r_i one-dimensional reflection vector fields, and $q_i(t')$ a process that only changes on the boundary for reflection. Here we use inward normal reflection on the boundary of the domain.

We can rewrite (6) above compactly as

$$d\boldsymbol{u} = \boldsymbol{\nabla}_u \Phi_{\boldsymbol{x}} dt' + \nu d\boldsymbol{w}(t') + \boldsymbol{r}(\boldsymbol{u}) d\boldsymbol{q}(t'), \tag{7}$$

where

- the dynamics are in the parameter time scale variable t',
- $\boldsymbol{w}$ is a vector of zero-mean, unit-variance normal random variables from a Wiener process,
- $\boldsymbol{r}$ a matrix of one-dimensional reflection vector fields coordinate functions, and
- $\boldsymbol{q}(t')$ a process that only changes on the boundary of the parameter domain for reflection.

Below, we develop the stochastic dynamical system that yields the Gibbs measure as the stationary measure. See example 3.8 in [15] and previous arguments for a rigorous proof, and also [10, 11, 13, 15, 8, 16] for stochastic/reflection and analytic techniques for semigroups and convergence rate bounds.

Proposition 1. *Let $\rho(\boldsymbol{u})$ be the joint density function over parameter domain Ω_u for a fast-tuning control problem (as outlined above) with N control variables. Let $\Phi_{\boldsymbol{x}}$ be the objective function with fast and slow timescales for parameters and dynamical variables, respectively (as in* (5) *above). Consider the dynamics*

$$d\boldsymbol{u} = \boldsymbol{\nabla}_u \Phi_{\boldsymbol{x}} dt' + \nu d\boldsymbol{w}(t') + \boldsymbol{r}(\boldsymbol{u}) d\boldsymbol{q}(t') \tag{8}$$

[2] "$dw_i(t')$" is often referred to as a Gaussian white noise.

where $\mathbf{u} \in \Omega_u$*, and* $\mathbf{w}$ *is a vector of* N *standard Wiener processes which are identical and independent across agents and time. Furthermore, the* w_i *have mean zero and variance one, and reflecting boundary conditions* [3] *are used.*

If the process $\mathbf{u}(t')$ *satisfies the dynamics of* (8)*, then the stationary density satisfies the Fokker-Planck equation*

$$\frac{\partial \rho(\boldsymbol{u}, t')}{\partial t'} = 0 = -\boldsymbol{\nabla} \cdot [\boldsymbol{\nabla} \Phi_{\boldsymbol{x}}(\boldsymbol{u}(t'))\rho(\boldsymbol{u}, t')] + \tfrac{\nu^2}{2} \nabla^2 \rho(\boldsymbol{u}, t') \tag{9}$$

and hence the stationary measure for variance ν^2 *is the Gibbs state* [4] *(*[15] *corollary 1 and example 3.8)*

$$\rho(\boldsymbol{u}, t') = \rho(\boldsymbol{u}) = \frac{\exp\left(\frac{2}{\nu^2}\Phi_{\boldsymbol{x}}(\boldsymbol{u})\right)}{\int_{\Omega_u} \exp\left(\frac{2}{\nu^2}\Phi_{\boldsymbol{x}}(\boldsymbol{u'})\right) d\boldsymbol{u'}}. \tag{10}$$

In statistical mechanics, the term in the exponent of (10) is $-E(\boldsymbol{u})/(kT)$, where k is Boltzmann's constant, T is temperature, and $E(\boldsymbol{u})$ is the energy of configuration $\boldsymbol{u}$. Hence the analogy of a fast-tuning parameter control theory problem to statistical mechanics is that ν^2 (deviation from rationality; influence of the noise in dynamics (6)) is proportional to 'temperature' and the objective function Φ (we wish to *maximize*) is the negative 'energy' of the system [1].

An alternate approach that yields the Gibbs measure is a maximum information-entropy solution where the objective function (5) is constrained to have a fixed mean (see [1]). If the objective function is constrained to have a mean below its maximum, then this constraint is enforcing the degree of bounded rationality. This is because perfectly rational agents would choose the maximum of the objective function. In the constrained maximum entropy view, where the mean of the objective function is constrained, agents play other strategies and "arbitrage" information out of the system by exploring other strategies to try to achieve a greater maximum of the objective function.

[3] We use a normal reflection vector field on the boundary $\partial\Omega_u$ as in [15].

[4] We use the convenient notation $d\boldsymbol{u} = \prod_{i=1}^{N} du_i$

A fluctuation-dissipation argument then relates the constrained maximum information entropy solution to the stochastic dynamical solution (both Gibbs measures) and the "temperature" T is seen to be proportional to the square of the fluctuation parameter ν^2 by

$$T := 1/\beta = \frac{\nu^2}{2}. \tag{11}$$

This results in the measure

$$\rho_{\boldsymbol{x},\beta}(\boldsymbol{u}) := \frac{\exp\left(\beta \Phi_{\boldsymbol{x}}(\boldsymbol{u})\right)}{\int_{\Omega_u} \exp\left(\beta \Phi_{\boldsymbol{x}}(\boldsymbol{u}')\right) d\boldsymbol{u}'}, \tag{12}$$

Let $C(\Omega_u)$ be the set of continuous function on the parameter domain. An **observable** $g \in C(\Omega_u)$ is a continuous function on pure state space such as the value of a specific parameter u_k, a variance-type variable $(u_k - \mu)^2$ for a constant μ, etc. We denote the expected value of an observable g with respect to the Gibbs state (12) by

$$E_{\boldsymbol{x},\beta}[g] := \int_{\Omega_u} g(\boldsymbol{u}) \rho_{\boldsymbol{x},\beta}(\boldsymbol{u}) d\boldsymbol{u}. \tag{13}$$

Finally, annealing the parameters implies that each possible direction $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u})$ the ambient dynamical curve will evolve from point $\boldsymbol{x}$ is weighted by the Gibbs measure. We take a weighted sum over all possible directions, since there is a probability every direction is chosen, to get

$$d\boldsymbol{x} = E_{\boldsymbol{x},\beta}\left[\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u})\right] dt. \tag{14}$$

The control theory problem is then reduced to solving an annealed version of (3):

$$\boldsymbol{x}(t) = \boldsymbol{x}(0) + \int_0^t E_{\boldsymbol{x}(\tau),\beta(\tau)}\left[\boldsymbol{f}(\boldsymbol{x}(\tau), \boldsymbol{u})\right] d\tau, \qquad \boldsymbol{x}(0) = \boldsymbol{x}_0, \;\; t \in [0, T], \tag{15}$$

where we point out that the inverse temperature β, which is proportional to the temporal variance of noise (i.e., fluctuations from rational decisions), is a function of the ambient system time t of the state variables $\boldsymbol{x}(t)$. For example, if agents gain more information about the system over time, the "temperature" of the system will decrease (i.e., optimal parameters will have a higher probability of being chosen).

In the case of a two-market speculative and hedging model [4, 5, 6], it was shown that a phase transition can occur. This results in two possible infinite-agent Gibbs measures at lower temperatures, $\rho^-_{\boldsymbol{x},\beta}$ and $\rho^+_{\boldsymbol{x},\beta}$, which would bifurcate the dynamics in (15). We elaborate on this as an example in a subsequent section.

3. Example: Linear Quadratic Controller

A linear quadratic contoller has quadratic operational cost and is constrained to move on a linear manifold. Let $\mathbb{M}^{m\times n}$ denote the set of m-by-n matrices.

An example of a classical control problem is given by the ambient dynamics

$$\frac{d}{dt}\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}\boldsymbol{\alpha}(t), \qquad \boldsymbol{x}(0) = \boldsymbol{x}_0,\ \ t \in [0, T], \tag{16}$$

for $1 \leq i \leq M$ and matrices $\boldsymbol{A} \in \mathbb{M}^{M\times M}$ and $\boldsymbol{B} \in \mathbb{M}^{M\times N}$. The objective function that is to be *minimized* is

$$\tilde{\Phi}(\boldsymbol{\alpha}) \ = \int_0^T [\boldsymbol{x}(t)^*\boldsymbol{P}\boldsymbol{x}(t) + \boldsymbol{\alpha}(t)^*\boldsymbol{Q}\boldsymbol{\alpha}(t) + 2\boldsymbol{x}(t)^*\boldsymbol{R}\boldsymbol{\alpha}(t)]\, dt + \boldsymbol{x}(T)^*\boldsymbol{S}\boldsymbol{x}(T), \tag{17}$$

for the matrix $\boldsymbol{A}$ and symmetric non-negative definite matrices $\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{S} \in \mathbb{M}^{M\times M}$, $\boldsymbol{R} \in \mathbb{M}^{M\times N}$, with $\boldsymbol{x}^*$ representing the Hermitian transpose of the vector $\boldsymbol{x}$.

Here, we consider the case $\boldsymbol{R} = \boldsymbol{0}$ in (17) (an example of the non-zero case is in Section 4). The fast-tuning, boundedly-rational objective function as in (5) is to *maximize*

$$\Phi_{\boldsymbol{x}(t)}(\boldsymbol{u}) \ = -\boldsymbol{x}(t)^*\boldsymbol{P}\boldsymbol{x}(t) - \boldsymbol{u}^*\boldsymbol{Q}\boldsymbol{u}. \tag{18}$$

In this case, the state variables separate from the control variables in the Gibbs measure in (12) to yield

$$\begin{aligned} \rho_\beta(\boldsymbol{u}) &:= \frac{\exp\left(-\beta\left[\boldsymbol{x}(t)^*\boldsymbol{P}\boldsymbol{x}(t) + \boldsymbol{u}^*\boldsymbol{Q}\boldsymbol{u}\right]\right)}{\int_{\Omega_u} \exp\left(-\beta\left[\boldsymbol{x}(t)^*\boldsymbol{P}\boldsymbol{x}(t) + \boldsymbol{u}'^*\boldsymbol{Q}\boldsymbol{u}'\right]\right) d\boldsymbol{u}'} \\ &= \frac{\exp\left(-\beta\boldsymbol{u}^*\boldsymbol{Q}\boldsymbol{u}\right)}{\int_{\Omega_u} \exp\left(-\beta\boldsymbol{u}'^*\boldsymbol{Q}\boldsymbol{u}'\right) d\boldsymbol{u}'}, \end{aligned} \tag{19}$$

noting that the dependence of ρ_β on $\boldsymbol{x}(t)$ was removed. The annealed dynamics as in (14) become

$$\begin{aligned} d\boldsymbol{x} &= E_\beta\left[\boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{u}\right] dt \\ &= \left(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}E_\beta\left[\boldsymbol{u}\right]\right) dt \\ &= \left(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{z}(\beta)\right) dt \end{aligned} \tag{20}$$

where $\boldsymbol{z}$ is a vector function that depends explicitly on β, and implicity on (ambient system dynamical) time t since β can vary with time. We point out that $\boldsymbol{z}$ does not depend on $\boldsymbol{x}$ in this case, and that lower values of the inverse-temperature parameter $\beta = 1/T$ correspond to less deviations from optimality.

4. Example: Speculative and Hedging Economic Model

A boundedly-rational speculative and hedging game-theoretic model with a long-term investor ("Air") and a finite number N of short-term "Bank" investing speculators was investigated in [4, 5, 6, 7]. All of these Bank players are labeled by elements in the set $\Lambda_N = \{1, 2, \ldots, N\}$. At any moment in time, a Bank player $i \in \Lambda_N$ can select an action or strategy $(y_1^{(i)}, y_2^{(i)}) \in F$ and the $y_1^{(i)}$ and $y_2^{(i)}$ are the **strategy variables**, which are the **parameters** of the control problem. The amount that Air invests will be the **state variable** of the control problem.

The strategy $y_1^{(i)}$ represents the proportion of its resources that Bank i spends on the oil spot market, and $y_2^{(i)}$ represents the proportion of its resources spent on the (US) dollar futures market from its total resources $M > 0$. A **parameter configuration** $\mathbf{y}$ is a point in parameter space, and describes a state of the speculators:

$$\mathbf{y} = \left((y_1^{(1)}, y_2^{(1)}), (y_1^{(2)}, y_2^{(2)}), \ldots, (y_1^{(N)}, y_2^{(N)})\right), \tag{21}$$

where each $(y_1^{(i)}, y_2^{(i)}) \in F$. The set of all possible *parameter* configurations of the game is

$$\Omega_{y,\Lambda_N} \coloneqq \prod_{i \in \Lambda_N} F^{(i)}, \tag{22}$$

which is called (pure) **parameter space**. The $F^{(i)} := F$ here is the diamond set

$$F := \left\{ \|(y_1^{(i)}, y_2^{(i)})\|_1 = \left|y_1^{(i)}\right| + \left|y_2^{(i)}\right| \leq 1 \right\}. \tag{23}$$

The long-term investing Air is assigned zero as its player number. Its strategy variable is the *state variable*

$$x \in [0, 1] \tag{24}$$

which represents the proportion of its resources spent on purchasing oil futures as a hedge. The remaining proportion of Air's resources, $1 - x$, is spent purchasing jet fuel on the spot market. Air will try to maximize their payoff function from their investment in the oil market, and thus follow the direction of maximum payoff (we'll assume this in a purely rational manner):

$$dx = \frac{\partial}{\partial x} f_O^{(0)}(x, \mathbf{y}) dt \tag{25}$$

The **payoff function for Bank** i, $1 \leq i \leq N$, below is the sum of the payoffs from the oil and dollar markets:

$$\begin{aligned} f^{(i)}(x, \mathbf{y}) &= f_O^{(i)} + f_\$^{(i)} \\ &= y_1^{(i)} \sum_{j=1}^{N} \left(\tfrac{E}{N} y_1^{(j)} - \tfrac{D}{N} y_2^{(j)} \right) \\ &\quad + y_2^{(i)} \sum_{j=1}^{N} \left(-\tfrac{K}{N} y_1^{(j)} + \tfrac{J}{N} y_2^{(j)} \right) - hxy_2^{(i)}. \end{aligned} \tag{26}$$

Each speculating Bank agent i could follow the gradient of their individual payoff function f_i to attempt to maximize profit. But in the case that $D = K$ in (26) above, the condition for a *potential game* [17] is satisfied, and there is a function V that coordinates all of the individual payoff functions into the single function V as a result of the condition (for $\alpha = 1, 2$)

$$\tfrac{\partial}{\partial y_\alpha^{(i)}} f^{(i)}(\mathbf{y}) = \tfrac{\partial}{\partial y_\alpha^{(i)}} V(\mathbf{y}). \tag{27}$$

Local maximums of the potential V correspond to Nash equilibriums of the game, and actually refine the Nash equilibriums [9]. Players would have no incentive to deviate from an absolute maximum of the potential on the compact set Ω_{y,Λ_N}, since they could not increase their profit above the value at an absolute maximum.

After the change of parameter variables for $1 \leq i \leq N$,

$$u_1^{(i)} = \left(y_2^{(i)} + y_1^{(i)}\right)/2, \qquad u_2^{(i)} = \left(y_2^{(i)} - y_1^{(i)}\right)/2, \tag{28}$$

$u_\alpha^{(i)} \in [-1/2, 1/2]$ for $\alpha = 1, 2$, the rotated domain

$$\tilde{F}^{(i)} := [-1/2, 1/2] \bigtimes [-1/2, 1/2] \tag{29}$$

is a square, and the rotated configuration space is

$$\Omega_{u,\Lambda_N} := \prod_{i \in \Lambda_N} \tilde{F}^{(i)} = [-1/2, 1/2]^{2N}. \tag{30}$$

The potential is then

$$\begin{aligned} V(x, \mathbf{u}) = \sum_{i,j=1}^{N} \tfrac{I_-}{N} u_1^{(i)} u_1^{(j)} + \sum_{i,j=1}^{N} \tfrac{I_+}{N} u_2^{(i)} u_2^{(j)} - \tfrac{2I}{N} \sum_{i,j=1}^{N} u_1^{(i)} u_2^{(j)} - \tfrac{2I}{N} \sum_{i=1}^{N} u_1^{(i)} u_2^{(i)} \\ + \tfrac{I_-}{N} \sum_{i=1}^{N} \left[u_1^{(i)}\right]^2 + \tfrac{I_+}{N} \sum_{i=1}^{N} \left[u_2^{(i)}\right]^2 + hx \sum_{i=1}^{N} \left(u_1^{(i)} + u_2^{(i)}\right) \end{aligned} \tag{31}$$

where

$$I := (J - E)/2, \tag{32}$$

$$I_+ := (J + E)/2 + K, \tag{33}$$

$$I_- := (J + E)/2 - K. \tag{34}$$

Let $\mathbf{1}$ be a $2N$-dimensional vector with all components equal to one. It was shown in [5] (c.f., section 2) that the potential in (31) can be written as a quadratic form

$$V(x, \mathbf{u}) = \mathbf{u}^* Q \mathbf{u} + hx \mathbf{1} \cdot \mathbf{u}, \tag{35}$$

which is a fast-tuning objective function that fast-trading speculators wish to maximize. We point out that Q is not necessarily non-negative definite. Note that the state variable x is coupled to the parameters $\mathbf{u}$, and x can be considered constant as speculators reach a stationary state, since Air's investment in the oil market is over a relatively long term and changes on a much longer time scale than the trading of speculators.

The boundedly-rational speculators will then follow the stochastic dynamics

$$d\mathbf{u} = \boldsymbol{\nabla}_{\mathbf{u}} V dt' + \nu d\mathbf{w}(t') + \mathbf{r}(\mathbf{u}) d\boldsymbol{q}(t'), \tag{36}$$

with $\mathbf{u} = \left(u_1^{(1)}, u_2^{(1)}, \ldots, u_1^{(N)}, u_2^{(N)}\right)$, $d\mathbf{w} = \left(dw_1^{(1)}, dw_2^{(1)}, \ldots, dw_1^{(N)}, dw_2^{(N)}\right)$, $d\boldsymbol{q} = \left(dq_1^{(1)}, dq_2^{(1)}, \ldots, dq_1^{(N)}, dq_2^{(N)}\right)$ as outlined in section Section 2.

The stationary state of (36) is the Gibbs measure

$$\rho_{x,\beta}(\mathbf{u}) \coloneqq \frac{\exp\left(\beta V(x, \mathbf{u})\right)}{\int_{\Omega_{u,\Lambda_N}} \exp\left(\beta V(x, \mathbf{u}')\right) d\mathbf{u}'}. \tag{37}$$

For a finite number of speculators, the stationary state is unique, and we have exactly the fast-tuning boundedly rational control system as outlined above. But as is common in economics, the case of perfect competition $N \to \infty$ is considered, in which there is a statistically very large number of speculators. In this case, it was shown in [6] that two stationary states are possible for temperatures T below a critical temperature T_{crit}, ρ_β^+ and ρ_β^-. This gives rise to a bifurcation in the ambient state dynamics of Air, which formally is:

$$dx = \frac{d}{dx} E_\beta^- \left[f_O^{(0)}(x, \mathbf{u}) \right] dt \tag{38}$$

and

$$dx = \frac{d}{dx} E_\beta^+ \left[f_O^{(0)}(x, \mathbf{u}) \right] dt. \tag{39}$$

5. Conclusion

We have presented a new and systematic way to incorporate fast-switching parameters into control theory with the more realistic assumption of bounded rationality -

a real, physical control mechanism for a complex system will not have all the information for the system that will yield a truly optimal control. With the ever increasing speed of computers, the time it takes to adjust parameters is much smaller than could ever have been imagined in the past, when control theory was initiated. Applications have already been implemented in a private sector setting, involving a control system that works with large scale building and power grid systems: this is currently under development at a company at which the author has done work. A much more theoretical application can be found in economics [3], where fast-trading speculators are effective "parameters" of a system, and the long-term investor is the "control" aspect of the system.

It is novel and pleasing that statistical mechanics turns out to integrate into control theory to extend it to the scenario of fast-switching parameters. Further research will be to work out the case of a control system with both slow and fast switching parameters.

References

[1] M. Campbell, A Gibbsian Approach to Potential Game Theory, draft at http://arxiv.org/abs/cond-mat/0502112, 2005

[2] M. Campbell, Inevitability of Collusion in a Coopetitive Bounded Rational Cournot Model with Increasing Demand, Journal of Mathematics Economics and Finance, 2016, 2 issue 1, p.7-19, http://dx.doi.org/10.14505/jmef.v2.1(2).01

[3] M. Campbell, Speculative and Hedging Interaction Model in Oil and U.S. Dollar Markets - Long Term Investor Dynamics and Phases, submitted, 2019

[4] D. Carfi and M. Campbell, Bounded Rational Speculative and Hedging Interaction Model in Oil and U.S. Dollar Markets, Journal of Mathematics Economics and Finance, 2015, 1 issue 1, p.29-47, http://dx.doi.org/10.14505/jmef.01

[5] D. Carfi and M. Campbell, Bounded Rational Speculative and Hedging Interaction Model in Oil and U.S. Dollar Markets: Nash Equilibrium, Atti della Accademia Peloritana dei Pericolanti Journal, 2018, Vol. 96, No. S1, A1

[6] M. Campbell and D. Carfi, Speculative and Hedging Interactions Model in Oil and U.S. Dollar Markets: Phase Transition, Journal of Statistical Physics, 2018, vol. 170, issue 1, p. 165-205, https://link.springer.com/article/10.1007%2Fs10955-017-1915-7

[7] D. Carfi and F. Musolino, Speculative and Hedging Interaction Model in Oil and U.S. Dollar Markets with Financial Transaction Taxes, Economic Modelling, 2014, 37 no. 0, p.306-319, http://dx.doi.org/10.1016/j.econmod.2013.11.003

[8] J. A. Carillo, A. Jüngel, P. Markovich, G. Toscani, A. Unterreiter, Entropy dissipation methods for degenerate parabolic problems and generalized Sobolev inequalities, Monatshefte für Mathematik, 2001 April, **133** (1), p.1-82

[9] O. Carbonell-Nicolau and R. McLean, Refinements of Nash equilibrium in potential games, Theor. Econ., 2014, 9, p.555-582

[10] J. G. Dai and R. J. Williams, Existence and uniqueness of semimartingale reflecting Brownian motions in convex polyhedrons, Theory of Probability and Its Applications, 1995, **40**, p.1-40, Correctional note, 50 (2006), 346-347.

[11] P. Dupuis and R. J. Williams, Lyapunov functions for semimartingale reflecting Brownian motions, Annals of Probability, 1994, **22**, p.680-702

[12] L. Evans, An Introduction to Mathematical Optimal Control Theory, version 0.2, open source

[13] J. M. Harrison and R. J. Williams, Multidimensional reflected Brownian motions having exponential stationary distributions, Annals of Probability, 1987, **15**, p.115-137

[14] P. Kumar, W. Kohn, and Z. Zabinsky, Near Optimal Hamiltonian-Control and Learning via Chattering, https://arxiv.org/abs/1703.06485, 2017

[15] W. Kang and K. Ramanan, Characterization of Stationary Distributions of Reflected Diffusions, Annals of Applied Probability, 2014, Vol. 24, No. 4, p.1329-1374

[16] S. A. Mohammed and M. K. R. Scheutzow, The stable manifold theorem for non-linear systems with memory: I. Existence of the semiflow, Journal of Functional Analysis, 2003, **205**, p.271-305

[17] D. Monderer and L. S. Shapley, Potential Games, Games and Economic Behavior, 1996, 14, p.124-143