Tadeusz Banek

Edward Kozłowski

LQG MONOTONE FOLLOWER MODEL OF CHANGE CONTROL IN TURBULENT ENVIRONMENT

Abstract

Change control (ChC) and monotone LQG follower (MFP) are two problems considered in this paper. Analogies and similarities between them allow for introducing a common parametrization, which can be used for translation of results obtained for the follower problem into the jargon of change control. Analogies and similarities between these problems are recognized on two levels: movement's description and criteria of control. That is the first moment when multicriteria approach is taken into account. On the other hand, there are many possible criteria which can, and should, be taken into account in deciding how to apply results coming from solutions of MFP into actions for ChC. That is the second moment for multicriteria approach to work.

Keywords

Stochastic control, change control, monotone follower, LQG problems, incremental value of information, Lagrange multipliers.

Introduction

The change control problem (ChC) has been by now extensively studied in literature (see [3], [4], [5], [6], [10], [11], [12], [14], [15], [16], [21] for example). These studies are qualitative rather then quantitative and till now there is lack of mathematical methods which are necessary for the creation of a general and coherent methodology. This paper is a step in this direction. We take a multicriterial approach as a cornerstone and use it in several aspects. It is particularly convenient when analogies and/or similarities are discussed, because they are, by definition, dependent upon criteria selected. The paper is organized as follows: in the next section we present formulations of two problems, the Linear Quadratic Gaussian (LQG) follower problem (FP) in R^n and the change control problem. From a general cybernetic/behavioral perspective it can be noticed that the problems experience analogies or similarities and create more complex relationships than just a set of random events. Analogies and similarities between these problems are recognized on two levels: movement's description and criteria of control. That is the first moment when multicriterial approach is taken into account. On the other hand, there are many possible criteria which can, and should, be taken into account in deciding how to apply results coming from solutions of MFP into actions for ChC. That is the second moment for multicriterial approach to be taken into account. This is our starting point for a new methodology explained in Subsection 1.2. We sketch here an idea. Mathematical results obtained for the FP in Subsections 2.2, 2.3, are in a quantitative form of optimal control formulae. The optimal control depends however on model parameters and selected criteria. Assuming that parameters and criteria in both problems are the same, optimal control laws can be characterized qualitatively using parametrization introduced in both problems. An optimal solution for the FP is – per analogy – recommended for the ChP. This methodology is applied in the LQG case in Subsection 2.3. Incremental value of information is introduced next and applied in the change control problem using the same approach. In the last section we offer a precise description of the essence of our methodology.

1. Problem Formulation

To make our presentation easier we begin by describing the Follower problem.

1.1. The Follower Problem in R^n

Let (Ω, F, P) be a complete probability space where the random variables $\xi_0, w_1, ..., w_N$, $\theta_1, ..., \theta_N$ are defined. They are assumed to be stochastically independent and such that

$$P(\theta_i = 0) = p = 1 - P(\theta_i = 1)$$

$$P(w_i \in A) = \int_A q(w)dw, \; P(\xi_0 \in A) = \int_A P(d\xi_0)$$

where $i = 1, ..., N$. For $f : R^n \times R^m \to R^n$, a measurable function called the *dynamic function*, let us define a stochastic system, called the Evader (E), via the iterative scheme

$$\xi_{i+1} = f(\xi_i, \varepsilon\theta_{i+1} w_{i+1}) \tag{1}$$

where $\varepsilon > 0$, and the product $\varepsilon\theta_{i+1} w_{i+1}$ models stochastic disturbances occurring in the system. Since $\theta_i = 1$ or 0, the disturbance εw_i occurs in time i or it does not. By allowing ε and p to have bigger or smaller values we can model the intensity of the random disturbances affecting the movement of E. Another system, called the Follower (F), is described by the iterative scheme

$$x_{i+1} = g(x_i, u_i) \tag{2}$$

where $g : R^n \times R^m \to R^n$ is a measurable function. Here, by u_i we denote the control action at the time i. As the Follower is allowed to know $\xi_0,...,\xi_i, x_0,...,x_i$ at the time i, only controls of the form

$$u_i = v_i(\xi_0,...,\xi_i, x_0,...,x_i) \tag{3}$$

where $v_i : R^{(i+1)n} \times R^{(i+1)n} \to R^m$ are Borel measurable functions, are admissible. For $h : R^n \times R^m \times R^n \to R_+$, a Borel measurable and bounded from below function, let us introduce a cost functional as the optimization criterion

$$J(u) = E\left[\sum_{j=0}^{N-1} h(x_i, u_i, \xi_{i+1})\right] \tag{4}$$

where E denotes expectation with respect to the measure P. The aim of the Follower is to find

$$\min_{u \in U} J(u) \tag{5}$$

where $U = \{u_i = v_i(\xi_0,...,\xi_i, x_0,...,x_i); i = 0,..., N-1\}$.

Example 1. *If* $a, b, c \geq 0$ *and*

$$h(x,u,\xi) = a\|g(x,u) - \xi\|^2 + b\|g(x,u) - x\|^2 + c\|u\|^2$$

then

$$J(u) = E\left[\sum_{j=0}^{N-1} a\|x_{j+1} - \xi_{j+1}\|^2 + b\|x_{j+1} - x_j\|^2 + c\|u_j\|^2\right]$$

(*where* $\|\cdot\|$ *denotes the Euclidean norm in* R^n) *describes a sum of penalties: the first due to the distance between the trajectories of E and F, the second due to the large jumps of F, and the third due to costs of controls. A Follower whose large jumps are costly* ($b >> a$, $b >> c$) *is called monotone.*

1.2. The Change Control Problem

The Change Manager (Project Manager, Steering Committee) has at least two options to steer his organization in the right direction and to put in place necessary and often unpopular changes. He may try to follow a leader (benchmarking), or force implementation of the virtual picture of his organization, an image, which he created as an optimal response for challenges coming from a turbulent environment. The picture, either virtual or real (benchmark), evolves in time in some unpredictable and random fashion. Even more, the intensity of changes in the organization's environment has a strong impact on the variability of pictures under considerations. Hence, the movement of the picture is very similar to those of the Evader from the previous section. However, the implementation of the necessary changes can be done in several different ways. For instance, the changes can be introduced step-by-step, implemented over time, allowing the organization to adapt gradually. Such policy aims at protecting the organization from unnecessary and often costly shocks, the effects of revolutionary changes. It is worth mentioning that too small, time consuming, prudently done, cosmetic or delayed changes are costly as well. The opportunity cost is a special name reserved for this cost in Economics. Hence, one may consider the MFP as a natural candidate to model the ChC problem. However, there are several drawbacks of such a model. Is the space R^n, including its elements, the proper object to model the state space of organizations? From the mathematical point of view, it certainly is not. It is true that some structures, DNA for instance, can be coded as a sequence of numbers. Nevertheless, it is not obvious that for any sequence of numbers one can find DNA which is coded by this sequence. Is any linear combination of structures a structure itself? Again, the answer cannot be affirmative. But structures are only part of a state space description of the organization under consideration. If linear combinations of the elements are not allowed, what kind of mathematical operations are? Perhaps the only thing we can estimate is the distance between the different states of organizations of the same kind or at least very similar. An open question is the completeness of this metric space, say (M,r), i.e., is any Cauchy sequence $m_j \in M$, $r(m_i,m_j) \to 0$, $i,j \to \infty$ convergent to $m_\infty \in M$? In terms of Management Sciences this question would be expressed as follows: is a process a making sequence of smaller (and smaller) changes always convergent to the result which can be identified as a state of the organization? Is it still the organization? If it was law respecting, civilized and honest – will it continue to be civilized, honest and the law respecting? If the answer is affirmative then M is a set of second Baire category, by the Baire-Hausdorff theorem (see [20], p. 11).

If it is negative then either case is possible. But even more important is that the Follower problem in (M,r) seems to be not mathematically tractable. This is demonstrated in the next sections. Given the above, only from the cybernetic perspective and assuming a high abstract level of reasoning, one may consider the MFP in R^n space as a model of the ChC problem. Quoting the great Stefan Banach at this point: "A mathematician is a person who can find analogies between theorems; a better mathematician is one who can see analogies between proofs and the best mathematician can notice analogies between analogies". We have discovered similarities and analogies between the problems. But what can one expect by solving the MFP instead of the ChC problem? Again, similarities and analogies between recommended policies. This is particularly important in the case of ChC. As a rule, such problems are difficult to model mathematically and obtaining quantitative solutions is, in general, hopeless. But the MFP in R^n is mathematically tractable. In simplest cases one can obtain explicit solutions in forms of formulas. These formulas depend explicitly upon problem parameters (for instance, the coefficients ε, p, a, b, c in the previous section) again. If the parameter space of the problem, say $\theta \in \Theta$, (the five-dimensional space of $\theta = (\varepsilon, p, a, b, c)$ in the example) is chosen, then for an optimal control $u^*(\theta)$, and a cost function $h(x, u^*(\theta), \xi)$, $\theta \in \Theta$, one can calculate the expected control energy $E[u^*(\theta)]^2$, the expected cost $E[h(x, u^*(\theta), \xi)]$, etc. Then the pairing $(\theta, u^*(\theta))$ clearly states how to act in ChC, if an element θ of Θ is a true parameter in the ChC Problem. Under these assumptions, quantitative results of the MFP can be characterized qualitatively and translated into the jargon of ChC. Such an approach is our methodological candidate for considering the ChC problems. In the next sections we are going to show how it works. But first we solve the MFP in some generality.

2. Solutions of the MFP

2.1. A linear-quadratic-gaussian case

To be more specific we shall apply the general results obtained so far to a LQG case. Assuming linearity in the right hand side of (1) and (2) we have

$$\xi_{i+1} = A\xi_i + \varepsilon\theta_{i+1}Bw_{i+1} \tag{6}$$

$$x_{i+1} = Cx_i + Du_i \tag{7}$$

where A, B, C, D are matrixes of appropriate dimensions. Denote the independence of random variables $x_1,...,x_n$ by $x_1 \perp ... \perp x_n$. Now we relax slightly the assumption imposed in Section 1.1. and introduce instead.

Condition 2. *Assume that:* (a) $\xi_0 \perp \theta_1 \perp ... \perp \theta_N$,
(b) $\theta_1 \perp ... \perp \theta_N \perp (w_1,...,w_N)$, ($c$) *the conditional distribution of* w_{i+1} *with respect to sub-sigma field* F_i *is* $N(G\xi_i, I)$, *where* G *is an* $m \times n$ *matrix, i.e., for any bounded, Borel measurable function* $h: R^n \to R$, *we have*

$$E\left[h(w_{i+1})|F_i\right] = (2\pi)^{-m/2} \int h(w) \exp\left[-\frac{1}{2}\|w - G\xi_i\|^2\right] dw. \tag{8}$$

Remark 3. *By conditioning the Evader's equation we get*

$$E\left[\xi_{i+1}|\xi_i\right] = A\xi_i + (1-p)BG\xi_i. \tag{9}$$

By the Condition (*2*) *we can model an effect of the inner resistances against changes in organizations* $(BG \approx -A)$, *an effect of approbation* $(BG \approx A)$, *or ambivalence* $(G = 0)$. *On the other hand, by putting*

$$C_1 = C + DH. \tag{10}$$

where H *is a* $m \times n$ *matrix, in place of* C *in* (*7*) *we have*

$$x_{i+1} = Cx_i + D[u_i + Hx_i]. \tag{11}$$

This shows that the substitution (*10*) *is a deterministic version of the Condition* (*2*). *Since* C *in* (*7*) *was arbitrary, we conclude that, to model the above effects, no additional condition for the Follower is needed.*

For a quadratic criterion as in the Example 1

$$J(u) = E\left[\sum_{j=0}^{N-1} a\|x_{j+1} - \xi_{j+1}\|^2 + b\|x_{j+1} - x_j\|^2 + c\|u_j\|^2\right] \tag{12}$$

we have the following

Proposition 4. *Assume:* (a) $a, b, c > 0$, (b) *Condition* (*2*),
(c) $\det\left[(a+b)D^T D + cI\right] \neq 0$. *Then*

$$u_j^* = \left[(a+b)D^T D + cI\right]^{-1} D^T \left\{a\left[A + (1-p)\varepsilon BG\right]\xi_j - \left[(a+b)C - bI\right]x_j\right\} \tag{13}$$

is the optimal control for the problem (*6*), (*7*), (*12*).

Proof. See appendix. ■

Conclusion 5. *In the case* $A = C,\ B = D$*, we get*

$$u_j^* = \left[(a+b)B^T B + cI\right]^{-1} B^T \left\{a\left[A + (1-p)\varepsilon BG\right]\xi_j - \left[(a+b)A - bI\right]x_j\right\} \quad (14)$$

Remark 6. u_j^* *is linear with respect to the state* (ξ_j, x_j)*, and system parameters* p, ε*.*

Remark 7. *Our results can be easily extended to cover nongaussian cases.*

Remark 8. *The crucial question is when a LQG Monotone Follower is a good model for Change Control. The most important assumption is linearity of state equations and normality of disturbances. From the equivalence theorem of J. Zabczyk* (*see* [*22*], *Chapter 3, Theorem 3.1.1, p. 26*) *it follows that without loss of generality one can choose as the space* (Ω, F, P)*, the basic probability space* $([0,1), B[0,1), \lambda_{[0,1)})$*, and as the noise, a sequence* $w_1, ..., w_N$ *of independent uniformly distributed random variables on* $[0,1)$*, or independent normally distributed random variables on* R*. Hence, it remains only to show linearity. Having a family* $F_{\xi_0,...,\xi_i}(x_0,...,x_i)$*,* $i = 0,...,N$*, of multidimensional distribution functions obtained from observations, one may apply the procedure described in the above-mentioned theorem of Zabczyk and as the result, we obtain a family of functions* $f_k(\xi, w)$*,* $k \in K$ *in* (*1*)*, such that the iterative scheme with this* $f_k(\xi, w)$ *as dynamic function will produce a sequence* $\xi_0, ..., \xi_N$ *with the distribution function equal to* $F_{\xi_0,...,\xi_i}(x_0,...,x_i)$*. By selecting members* $f_k(\xi, w)$ *of the set* $\{f_k(\xi, w) : k \in K\}$ *one may choose an element which is* (*1*) *linear, or* (*2*) "*as close as currently possible*" *to be linear. If the first case holds, or an approximation error* (*appropriately defined*) *in the second case is small, then we call a ChC Problem linear.*

2.2. Energy of optimal control

Denote $(a, b, c, p, G, \varepsilon) = \theta \in \Theta$. By introducing matrices

$$L = a\left[(a+b)D^T D + cI\right]^{-1} D^T \left[A + (1-p)\varepsilon BG\right] \equiv L(\theta) \quad (15)$$

$$H = \left[(a+b)D^T D + cI\right]^{-1} D^T \left((a+b)C - bI\right) \equiv H(\theta) \quad (16)$$

we can express the optimal control (13) and its energy in the form

$$u_j^* = L\xi_j + Hx_j \tag{17}$$

$$E_j^* = E\left[\left\|u_j^*\right\|^2 \middle| F_j\right] = x_j^T H^T H x_j + 2x_j^T H^T L\xi_j + \xi_j^T L^T L\xi_j \tag{18}$$

Remark 9. *Since the matrices* L, H *depend on the system parameters* $a, b, c, p, G, \varepsilon$*, so are the optimal control* u_j^* *and its energy* E_j^**. The functions*

$$\theta \in \Theta \to u_j^*(\theta) = L(\theta)\xi_j + H(\theta)x_j \tag{19}$$

$$\theta \in \Theta \to E_j^*(\theta) = x_j^T H^T(\theta)H(\theta)x_j + 2x_j^T H^T(\theta)L(\theta)\xi_j + \xi_j^T L^T(\theta)L(\theta)\xi_j \tag{20}$$

shows this dependence explicitly. Since E_j^* *is a measure of the effort at time j done by the Follower moving in* R^n*, hence – by analogy –* $\theta \to E_j^*(\theta)$ *shows this dependence on* θ *also for Change Control.*

In this section we are going to find a total control energy

$$W_0 = E\left[\sum_{j=0}^{N-1}\left\|u_j^*\right\|^2\right]. \tag{21}$$

Let us denote

$$W_N = 0 \tag{22}$$

$$W_j = E\left[\sum_{i=j}^{N-1}\left\|u_i^*\right\|^2 \middle| F_j\right] = \left\|u_j^*\right\|^2 + E\left[W_{j+1} \middle| F_j\right] \tag{23}$$

Proposition 10. Under the conditions of Proposition 4, we have

$$W_j = E\left[\sum_{i=j}^{N-1}\left\|u_i^*\right\|^2 \middle| F_j\right] = x_j^T Q_j x_j + x_j^T R_j \xi_j + \xi_j^T S_j \xi_j + K_j \tag{24}$$

for $j = 0,1,\dots,N-1$*, where*

$$Q_j = H^T H + (C + DH)^T Q_{j+1}(C + DH), \quad Q_{N-1} = H^T H \tag{25}$$

$$R_j = 2H^T L + 2(C + DH)^T Q_{j+1} DL + (C + DH)^T R_{j+1}\left[A + (1-p)\varepsilon BG\right] \tag{26}$$

$$R_{N-1} = 2H^T L \tag{27}$$

$$
\begin{aligned}
S_j &= L^T\left[I + D^T Q_{j+1} D\right]L + L^T D^T R_{j+1}\left[A + (1-p)\varepsilon BG\right] \\
&\quad + A^T S_{j+1}\left[A + 2(1-p)\varepsilon BG\right]
\end{aligned}
\tag{28}
$$

$$
S_{N-1} = 2L^T L \tag{29}
$$

$$
K_j = \sum_{i=j+1}^{N-1} (1-p)\varepsilon^2 tr\left(S_i BB^T\right), \quad K_{N-1} = 0 \tag{30}
$$

Proof. See appendix. ■

Conclusion 11. *The expected total energy of optimal control is given by*

$$
E\left[\sum_{i=0}^{N-1} \left\|u_i^*\right\|^2\right] = W_0 = x_0^T Q_0 x_0 + x_0^T R_0 \xi_0 + \xi_0^T S_0 \xi_0 + K_0 \tag{31}
$$

where Q_0, R_0, S_0, K_0 *are given in* (*25*)-(*30*).

Remark 12. *All remarks in the previous Remark are also valid for the expected total energy*

$$
\theta \in \Theta \to W_0(\theta) = x_0^T Q_0(\theta) x_0 + x_0^T R_0(\theta)\xi_0 + \xi_0^T S_0(\theta)\xi_0 + K_0(\theta) \tag{32}
$$

The functions $\theta \in \Theta \to Q_0(\theta), R_0(\theta), S_0(\theta), K_0(\theta)$ *are given in the previous Proposition. Dependences on* θ *are − by analogy − expected to hold for linear Change Control.*

2.3. Size of the jumps

We want to calculate the sum of the expected jumps

$$
V_0 = E\left[\sum_{i=0}^{N-1} \left\|x_{i+1} - x_i\right\|^2\right] \tag{33}
$$

and its dependence on the systems parameters. Since

$$
\left\|x_{i+1} - x_i\right\|^2 = \left\|(C - I)x_i + Du_i^*\right\|^2 = \left\|(C + DH - I)x_i + DL\xi_i\right\|^2 \tag{34}
$$

hence

$$
V_0 = E\left[\sum_{i=0}^{N-1} \left\|(C + DH - I)x_i + DL\xi_i\right\|^2\right] \tag{35}
$$

Dennote

$$V_N = 0 \tag{36}$$

$$V_j = E\left[\sum_{i=j}^{N-1} \left\|(C + DH - I)x_i + DL\xi_i\right\|^2 \middle| F_j\right] = \left\|x_{j+1} - x_j\right\|^2 + E\left[V_{j+1} \middle| F_j\right] \tag{37}$$

Proposition 13. *Under the conditions of Proposition 4, we have*

$$\begin{aligned} V_j &= E\left[\sum_{i=j}^{N-1} \left\|(C + DH - I)x_i + DL\xi_i\right\|^2 \middle| F_j\right] \\ &= x_j^T \widetilde{Q}_j x_j + x_j^T \widetilde{R}_j \xi_j + \xi_j^T \widetilde{S}_j \xi_j + K_j \end{aligned} \tag{38}$$

for $j = 0,1,\ldots,N-1$*, where*

$$\widetilde{Q}_j = (C + DH - I)^T (C + DH - I) + (C + DH)\widetilde{Q}_{j+1}(C + DH)$$

$$\widetilde{Q}_{N-1} = (C + DH - I)^T (C + DH - I)$$

$$\widetilde{R}_j = 2(C + DH - I)^T DL + 2(C + DH)\widetilde{Q}_{j+1} DL + (C + DH)^T \widetilde{R}_{j+1}\left[A + (1-p)\varepsilon BG\right]$$

$$\widetilde{R}_{N-1} = 2(C + DH - I)^T DL$$

$$\widetilde{S}_j = L^T D^T \left[I + \widetilde{Q}_{j+1}\right] DL + L^T D^T \widetilde{R}_{j+1}\left[A + (1-p)\varepsilon BG\right] + A^T \widetilde{S}_{j+1}\left[A + 2(1-p)\varepsilon BG\right]$$

$$\widetilde{S}_{N-1} = L^T D^T DL$$

$$K_j = \sum_{i=j+1}^{N-1} (1-p)\varepsilon^2 tr\left(S_i BB^T\right), \quad K_{N-1} = 0 \tag{39}$$

Proof. See appendix. ■

Conclusion 14. *The expected sum of the jumps is given by the formula.*

$$E\left[\sum_{i=0}^{N-1} \left\|x_{i+1} - x_i\right\|^2\right] = V_0 = x_0^T \widetilde{Q}_0 x_0 + x_0^T \widetilde{R}_0 \xi_0 + \xi_0^T \widetilde{S}_0 \xi_0 + K_0 . \tag{40}$$

Since the matrices H, L are functions of θ, so are the matrices $\widetilde{Q}_0, \widetilde{R}_0, \widetilde{S}_0$ and the scalar K_0. From the Proposition above we get formulas for the expected value of individual jumps and their sum. The functions

$$\theta \in \Theta \to \varphi_{x,y}(\theta) = \left\|(C + DH(\theta) - I)x + DL(\theta)y\right\|^2$$

$$\theta \in \Theta \to \Phi_{x,y}(\theta) = x^T \widetilde{Q}_0(\theta)x_0 + x^T \widetilde{R}_0(\theta)y + y^T \widetilde{S}_0(\theta)y + K_0(\theta)$$

define explicitly dependences of individual jumps sizes and their sum on system's parameters.

In order to answer the question how the recommended changes could or should depend upon the turbulences in the organization's environment, we have to restrict the domain Θ of $\varphi_{x,y}(\theta)$ and $\Phi_{x,y}(\theta)$. Let's fix the coordinates (a,b,c,G) of $\theta=(a,b,c,p,G,\varepsilon)$ and let $(p,\varepsilon)\in[0,1]\times R_+$ be free. Denote

$$\phi_{x,y}^{a,b,c,\alpha}(\cdot,\cdot)=\varphi_{x,y}(a,b,c,\cdot,\alpha,\cdot) \tag{41}$$

$$\Psi_{x,y}^{a,b,c,\alpha}(\cdot,\cdot)=\Phi_{x,y}(a,b,c,\cdot,\alpha,\cdot) \tag{42}$$

Hence (41), (42) are restrictions of $\varphi_{x,y}(\theta)$, $\Phi_{x,y}(\theta)$ on $[0,1]\times R_+$ and are the functions of interest. Hence, by setting

$$\chi=(C+DH-I)x+aD\left[(a+b)D^TD+cI\right]^{-1}D^TAy\,,$$

$$z=aD\left[(a+b)D^TD+cI\right]^{-1}D^TBGy\,,$$

$$\rho=(p-1)\varepsilon$$

we obtain, after some transformations,

$$\phi_{x,y}^{a,b,c,\alpha}(p,\varepsilon)=\|\chi-\rho z\|^2\,.$$

Conclusion 15. *The expected conditional jumps is a quadratic function* $\rho\in R\to\|\chi-\rho z\|^2$ *of the turbulence parameter* $\rho=(p-1)\varepsilon$. *Its minimum* $\left\|\chi-\frac{\chi^Tz}{\|z\|^2}z\right\|^2$ *is achieved at* $\rho=\frac{\chi^Tz}{\|z\|^2}$. *Certainly, this minimum also depends on the remaining system parameters, namely* a,b,c,G. *According to our cybernetic/behavioral perspective we claim that the relations describing the LQG Monotone Follower Problem solution given in this section are also valid for the LQG Change Control* (*see Remark 12*).

3. Value of information about Evader's future movement

In any follower problem any information about the evader's future movements can change the pursuer's strategy radically. One may choose, for instance, to aim at the virtual point of intersection of both trajectories, instead of using a pure follower's strategy, etc. Hence, any predictions coming from advisers and/or experts are valuable. But, what is the value? It is particularly interesting, how to price mathematically such information. In this section we introduce the notion of incremental value of information and apply it to the MFP.

Remark 16. *In stochastic optimization problems one can implement at least two approaches to defining the value of information. The first approach, leading to so called Incremental Value of Information* (*IVI*), *was initiated by M.H.A. Davis, M.A.H. Dempster, R.J. Elliott in* [*7*] *and by K. Back, S.R. Pliska in* [*1*] *and uses an idea of R.J.-B. Wets* [*19*]. *The second approach, initiated by T. Banek, R. Kulikowski in* [*2*] *and independently by M. Schweizer, D. Becherer in* [*18*], *is based on the idea that the information can be an object of trade and its value for a particular agent is a consequence of its utility. In this paper we follow the first method. By introducing the Lagrange multiplier we turn the optimization problem into a global minimization one over all controls from* $V \in L_p^m(N \times (\Omega, F, P))$ *which are* ξ^N *– measurable, i.e., take the form* $V = \{v_i(x^i, \xi^N), i = 0,1,\ldots,N-1\}$. *Secondly, the Lagrange multiplier can be interpreted as a price system for small violations of the constraint (the shadow price in* [*1*]), *in our case, small* ξ^N *– measurable perturbations of the controls. That approach is similar to small anticipative* (*allowed to know the future*) *perturbations considered in the paper* [*7*]. *Our price system may perhaps have some practical value for F who has an extra option, for instance;* (*1*) *he can predict the whole future movement of E, i.e.,* ξ^N (*or its part*) *by himself, doing, for instance, extensive* (*and costly*) *research, or* (*2*) *to buy a prognosis of* ξ^N *made by experts. The question of interest for F is: what is the right price for buying* ξ^N*? Our price system tells only how much costs a small violation of the constraint and thus can serve as a linear approximation.*

3.1. Subspace constrains and Lagrange multipliers

Let X be a Banach space with dual space X^*, and let S be a linear subspace of X. We define

$$S^{\perp} = \{x^* \in X^* :< x^*, x >= 0, \forall x \in S\}$$

where $< x^*, x >$ denotes the pairing between $x \in X$ and $x^* \in X^*$.

Let $\phi : X \to R$ be a Frechet differentiable functional and suppose that ϕ achieves its minimum over S at $x_0 \in S$. The Frechet derivative is a map $\phi' : X \to X^*$ such that for $h, x \in X$

$$\phi(x+h) = \phi(x) + \langle \phi'(x), h \rangle + o(\|h\|).$$

Lemma 17. *If ϕ achieves its minimum over S at $x_0 \in S$, then $\phi'(x_0) \in S^{\perp}$.*

Proof. If $\phi'(x_0) \notin S^{\perp}$ then there exists $h \in S$ such that $\langle \phi'(x_0), h \rangle = \delta > 0$. But then $\phi(x_0 - \varepsilon h) = \phi(x_0) - \varepsilon(\delta + o(\varepsilon)/\varepsilon)$ so that $\phi(x_0 - \varepsilon h) < \phi(x_0)$ for small ε.■

Theorem 18. *If $\phi : X \to R$ is Frechet differentiable and achieves its minimum over S at $x_0 \in S$, then there exists $\lambda \in S^{\perp}$ such that Lagrange functional*

$$L(x) = \phi(x) + \langle \lambda, x \rangle$$

is stationary at x_0, i.e., $L'(x_0) = 0$.

Proof. We have only to set $\lambda = -\phi'(x_0)$.■

3.2. Incremental value of information

To apply the above results to our problem, we take X to be the space $L_p^m(N \times (\Omega, F, P))$ of all controls $V = \{v_i(x^i, \xi^N), i = 0,1,\ldots, N-1\}$, i.e., anticipating controls which have access to information about the future movement of Evader F, and S to be a subspace of X of all controls

$U = \{u_i(x^i, \xi^i); i = 0,1,...,N-1\}$. It is clear that S is a linear subspace of X. Then X^* is $L_q^m(N \times (\Omega, F, P))$ space where $q = \frac{p}{p-1}$ and

$$S^{\perp} = \{\lambda \in X^* : E < \lambda, V >= 0, \forall V \in S\}.$$

The relationship between Gateaux and Frechet derivative of ϕ is that if the Gateaux derivative takes the form

$$E \sum \lambda_j v_j \tag{43}$$

for some $\lambda = col(\lambda_0, ..., \lambda_{N-1}) \in X^* = L_q^m(N \times (\Omega, F, P))$, then ϕ is Frechet differentiable and $\phi'(u) = \lambda$. Hence, from (43) we obtain

Theorem 19. *Assume that $\nabla_u h$ is bounded. Under the assumptions of Proposition 4*

$$\lambda_j = \nabla_u h(x_j, u_j^*, \xi_{j+1}) \tag{44}$$

for $j = 0,1,...,N-1$.

Proof. The RHS of (44) is bounded, hence it belongs to $L_q^m(N \times (\Omega, F, P))$ for any $q \geq 0$.■

As it was explained in the proof of Proposition 4, if u_j^* is an optimal control, then the equality

$$E\left[w_j^T \nabla_u h(x_j, u_j^*, \xi_{j+1})\right] = 0$$

must hold for any F_j – measurable function w_j. Comparing this with (43) we get (44). Since u_j^* is Markovian, i.e., it is of the form

$$u_j^* = v_j(x_j, \xi_j, \theta) \tag{45}$$

where θ is a parameter, $\theta \in \Theta$ and $v_j : R^n \times R^n \times \Theta \to R^m$ is some function, hence, from (1), (44), (45) we obtain

$$\lambda_j = \nabla_u h(x_j, v_j(x_j, \xi_j, \theta), f(\xi_j, \varepsilon \theta w_{j+1})) \tag{46}$$

This formula shows dependence of the IVI on $\theta \in \Theta$ explicitly.

3.2.1. A linear case

Taking into account that

$$h(x,u,\xi)=a\|Cx+Du-\xi\|^2+b\|Cx+Du-x\|^2+c\|u\|^2$$

we have

$$\begin{aligned}\frac{1}{2}\nabla_u h\left(x_j,u_j^*,\xi_{j+1}\right)&=aD^T\left(Cx_j+Du_j^*-\xi_{j+1}\right)+bD^T\left(Cx_j+Du_j^*-x_j\right)+cu_j^*\\&=D^T\left[(a+b)C-bI\right]x_j-aD^TA\xi_j+\left[(a+b)D^TD+cI\right]u_j^*-aD^T\varepsilon\theta_{j+1}Bw_{j+1}\end{aligned}\tag{47}$$

where u_j^* is given by (13).

Corollary 20. *Under the conditions of Propositon 4 we have*

$$\begin{aligned}\frac{1}{2}\lambda_j&=\left\{I-\left[(a+b)D^TD+cI\right]\left[(a+b)B^TB+cI\right]^{-1}\right\}D^T\left[(a+b)C-bI\right]x_j\\&+a\left\{\left[(a+b)D^TD+cI\right]\left[(a+b)B^TB+cI\right]^{-1}D^T\left[A+(1-p)\varepsilon aB\right]-D^TA\right\}\xi_j\\&-aD^T\varepsilon\theta_{j+1}Bw_{j+1}\end{aligned}\tag{48}$$

Proof. Substitution of (13) into (47) gives (48). Hence, it remains only to show that (50) belongs to $L_q^m\left(N\times(\Omega,F,P)\right)$ for any $q\geq 0$. Since x_j are deterministic, ξ_j and $\theta_{j+1}Bw_{j+1}$ independent, from (6) it follows that it is enough to have $\theta_{j+1}Bw_{j+1}\in L_q^m\left(N\times(\Omega,F,P)\right)$. But this is obvious, because θ_j, w_j are independent and Bernoulli, Gaussian distributed. ■

4. Essence

At the end we offer a precise description of the essence of our methodology.

Definition 21. *Let parameter space Θ can be decomposed into disjoint pieces $\Theta_j, j\in J$ a set of indexes, i.e., $\Theta=\bigcup_{j\in J}\Theta_j$, $\Theta_i\bigcap\Theta_j=0$, for $i\neq j$. We say that Θ shows an $F/\varepsilon,\sigma$ - distinction with respect to a family $\{\Theta_j : j\in J\}$, if there is a function $F:\Theta\to R^d, d\geq 1$, such that conditions*:

(*a*) $\quad\left\|F(\Theta_1)-F(\Theta_2)\right\|\leq\varepsilon,\quad \Theta_1,\Theta_2\in\Theta_j$ *for any* j

(b) $\quad \left\| F(\Theta_1) - F(\Theta_2) \right\| \geq 1/\delta, \quad \Theta_1 \in \Theta_i, \Theta_2 \in \Theta_j$ *for any* $\quad i \neq j$

hold for some $\varepsilon, \delta \geq 0$*. If* (a) *and* (b) *hold for any* $\varepsilon \geq 0$*, then we call it an* F/δ *– distinction.*

Remark 22. *If* Θ *shows an* $u^*/\varepsilon,\delta$ *– distinction* (u^*/δ *– distinction) with respect to* $\{\Theta_j : j \in J\}$*, where* u^* *is an optimal control for the MFP, then by our cybernetic/behavioral perspective, strategies* v^* *recommended for ChC should – per analogy – have* $v^*/\varepsilon,\delta$ *– distinction* (v^*/δ *– distinction*) *with respect to the same* $\{\Theta_j : j \in J\}$*. Similarly for energy* E^* *– distinction, size of the jumps* V *– distinction, etc.*

Remark 23. *If* Θ *shows an* $\lambda/\varepsilon,\delta$ *– distinction (*λ/δ *– distinction) with respect to* $\{\Theta_j : j \in J\}$*, where* λ *is a Lagrange multiplier for the MFP, then by our cybernetic/behavioral perspective, a Lagrange multiplier* v *for ChC should – per analogy – has* $v/\varepsilon,\delta$ *– distinction* (v^*/δ *– distinction*) *with respect to the same* $\{\Theta_j : j \in J\}$.

Remark 24. *It is possible to obtain in one ChC/MFP many recommendations with different* $\varepsilon,\delta's$*. It is quite obvious that any recommendation coming from the solutions of the MFP to ChC is as convincing as the value of* $\zeta = \min\left(1/\varepsilon, 1/\delta\right)$*. Hence, the Change Manager should select the recommendations properly and apply them or not, according to the order of* ζ *and his preferences coming from the multicriterial approach.*

Appendix

Proof of Proposition 4.

A standard way for obtaining an optimal control is to take weak variations of $J(u)$, i.e., ε – derivative of $J(u^* + \varepsilon v)$ at $\varepsilon = 0$, where u^* is the optimal control and v an element of U – a set of admissible controls (see [17] for example). The procedure gives the following equalities

$$E\left[aD^T\left(Cx_j + Du_j^* - \xi_{j+1}\right) + bD^T\left(Cx_j + Du_j^* - x_j\right) + cu_j^* \middle| F_j\right] =$$

$$\left[(a+b)D^T C - bD^T\right]x_j - aD^T E\left[\xi_{j+1}\middle|F_j\right] + \left[(a+b)D^T D + cI\right]u_j^* = 0 .$$

Thus

$$u_j^* = \left[(a+b)D^T D + cI\right]^{-1}\left(aD^T E\left[\xi_{j+1}\middle|F_j\right] + \left[bD^T - (a+b)D^T C\right]x_j\right). \tag{49}$$

Since w_i, θ_i are stochastically independent for $i = 1,\ldots,N$, we get

$$\begin{aligned} E\left[\xi_{j+1}\middle|F_j\right] &= \sum_{i=0}^{1} E\left[\xi_{j+1}\middle|F_j, \theta_{j+1} = i\right]P\left(\theta_{j+1} = i\right) \\ &= A\xi_j + (1-p)\varepsilon BG\xi_j = \left[A + (1-p)\varepsilon BG\right]\xi_j . \end{aligned} \tag{50}$$

Substitution of (50) into (49) gives (13). This shows that (13) is the right candidate for an optimal control. But $a,b,c > 0$ implies convexity and non--degeneracy of the problem, hence from the general results of the LQG theory (see R.S. Liptser, A.N. Shiryiaev [13], for instance) follows the sufficiency. ■

Proof of Proposition 10.

From (7) and (13) we have

$$x_{j+1} = Cx_j + Du_j^* = (C + DH)x_j + DL\xi_j . \tag{51}$$

Now, for $j = N-1$ we have

$$\begin{aligned} W_{N-1} &= E\left[\left\|u_{N-1}\right\|^2\middle|F_{N-1}\right] = x_{N-1}^T H^T H x_{N-1} + 2x_{N-1}^T H^T L\xi_{N-1} + \xi_{N-1}^T L^T L \quad \xi_{N-1} \\ &= x_{N-1}^T Q_{N-1} x_{N-1} + x_{N-1}^T R_{N-1}\xi_{N-1} + \xi_{N-1}^T S_{N-1}\xi_{N-1} . \end{aligned} \tag{52}$$

Assume that (26) is true for $j+1$. Then

$$\begin{aligned} W_j &= E\left[\sum_{i=j}^{N-1}\left\|u_i^*\right\|^2\middle|F_j\right] = \left\|u_j^*\right\|^2 + E\left[W_{j+1}\middle|F_j\right] \\ &= x_j^T H^T H x_j + 2x_j^T H^T L\xi_j + \xi_j^T L^T L\xi_j \\ &\quad + E\left[x_{j+1}^T Q_{j+1} x_{j+1} + x_{j+1}^T R_{j+}\xi_{j+1} + \xi_{j+1}^T S_{j+1}\xi_{j+1} + K_{j+1}\middle|F_j\right]. \end{aligned} \tag{53}$$

But from (6), (8) and the properties of conditional expectation we have

$$\begin{aligned} &E\left[x_{j+1}^T Q_{j+1} x_{j+1} + x_{j+1}^T R_{j+}\xi_{j+1} + \xi_{j+1}^T S_{j+1}\xi_{j+1} + K_{j+1} \middle| F_j\right] \\ &= E\left[\left[(C+DH)x_j + DL\xi_j\right]^T Q_{j+1}\left[(C+DH)x_j + DL\xi_j\right] \middle| F_j\right] \\ &+ E\left[\left[(C+DH)x_j + DL\xi_j\right]^T R_{j+}\xi_{j+1} + \xi_{j+1}^T S_{j+1}\xi_{j+1} + K_{j+1} \middle| F_j\right] \\ &= x_j^T (C+DH)^T Q_{j+1}(C+DH)x_j + 2x_j^T (C+DH)Q_{j+1}DL\xi_j \\ &+ \xi_j^T L^T D^T Q_{j+1} DL\xi_j + x_j^T (C+DH)^T R_{j+1}\left[A + (1-p)\varepsilon BG\right]\xi_j \\ &+ \xi_j^T L^T D^T R_{j+1}\left[A + (1-p)\varepsilon BG\right]\xi_j + \xi_j^T A^T S_{j+1} A\xi_j \\ &+ 2\varepsilon(1-p)\xi_j^T A^T S_{j+1} BG\xi_j + (1-p)\varepsilon^2 tr\left(S_{j+1}BB^T\right) + K_{j+1} \end{aligned} \tag{54}$$

Finally

$$\begin{aligned} W_j &= K_j + x_j^T\left[H^T H + (C+DH)^T Q_{j+1}(C+DH)\right]x_j \\ &+ x_j^T\left[2H^T L + 2(C+DH)Q_{j+1}DL + (C+DH)^T R_{j+1}\left[A + (1-p)\varepsilon BG\right]\right]\xi_j \\ &+ \xi_j^T\left[L^T\left[I + D^T Q_{j+1}D\right]L + L^T D^T R_{j+1}\left[A + (1-p)\varepsilon BG\right]\right]\xi_j \\ &\qquad + \xi_j^T\left[A^T S_{j+1}\left[A + 2(1-p)\varepsilon BG\right]\right]\xi_j \end{aligned} \tag{55}$$

what finish the proof. ■

Proof of Proposition 13.

For $j' N+1$, we have

$$\begin{aligned} V_{N-1} &= E\left[\left\|(C+DH-I)x_{N-1} + DL\xi_{N-1}\right\|^2 \middle| F_{N-1}\right] \\ &= x_{N-1}^T (C+DH-I)^T (C+DH-I)x_{N-1} \\ &+ 2x_{N-1}^T (C+DH-I)^T DL\xi_{N-1} + \xi_{N-1}^T L^T D^T DL\xi_{N-1} \\ &= x_{N-1}^T \widetilde{Q}_{N-1} x_{N-1} + x_{N-1}^T \widetilde{R}_{N-1}\xi_{N-1} + \xi_{N-1}^T \widetilde{S}_{N-1}\xi_{N-1} \end{aligned} \tag{56}$$

Assume that (40) is true for $j+1$. Then

$$\begin{aligned} V_j &= E\left[\sum_{i=j}^{N-1}\left\|(C+DH-I)x_i + DL\xi_i\right\|^2 \middle| F_j\right] = \left\|(C+DH-I)x_j + DL\xi_j\right\|^2 + E\left[V_{j+1}\middle| F_j\right] \\ &= x_j^T (C+DH-I)^T (C+DH-I)x_j + 2x_j^T (C+DH-I)^T DL\xi_j + \xi_j^T L^T D^T DL\xi_j \\ &+ E\left[x_{j+1}^T \widetilde{Q}_{j+1} x_{j+1} + x_{j+1}^T \widetilde{R}_{j+1}\xi_{j+1} + \xi_{j+1}^T \widetilde{S}_{j+1}\xi_{j+1} + K_{j+1} \middle| F_j\right]. \end{aligned} \tag{57}$$

Using (6), (8) and from conditional expectation properties, we obtain

$$
\begin{aligned}
& E\left[x_{j+1}^T \widetilde{Q}_{j+1} x_{j+1} + x_{j+1}^T \widetilde{R}_{j+1} \xi_{j+1} + \xi_{j+1}^T \widetilde{S}_{j+1} \xi_{j+1} + K_{j+1} \middle| F_j \right] \\
&= E\left[\left((C+DH)x_j + DL\xi_j\right)^T \widetilde{Q}_{j+1} \left((C+DH)x_j + DL\xi_j\right) \middle| F_j \right] \\
&\quad + E\left[\left((C+DH)x_j + DL\xi_j\right)^T \widetilde{R}_{j+1} \xi_{j+1} + \xi_{j+1}^T \widetilde{S}_{j+1} \xi_{j+1} + K_{j+1} \middle| F_j \right] \\
&= x_j^T (C+DH)^T \widetilde{Q}_{j+1} (C+DH) x_j + 2x_j^T (C+DH)^T \widetilde{Q}_{j+1} DL\xi_j + \xi_j^T L^T D^T \widetilde{Q}_{j+1} DL\xi_j \\
&\quad + x_j^T (C+DH)^T \widetilde{R}_{j+1} [A + (1-p)\varepsilon BG]\xi_j + \xi_j^T L^T D^T \widetilde{R}_{j+1} [A + (1-p)\varepsilon BG]\xi_j \\
&\quad + \xi_j^T A^T \widetilde{S}_{j+1} A\xi_j + 2\varepsilon(1-p)\xi_j^T A^T \widetilde{S}_{j+1} BG\xi_j + +(1-p)\varepsilon^2 tr\left(\widetilde{S}_{j+1} BB^T\right) + K_{j+1} .
\end{aligned}
\tag{58}
$$

Finally

$$
\begin{aligned}
V_j &= K_j + x_j^T \left[(C+DH-I)^T (C+DH-I) + (C+DH)^T \widetilde{Q}_{j+1} (C+DH)\right] x_j \\
&\quad + x_j^T \left[2(C+DH-I)^T DL + 2(C+DH)\widetilde{Q}_{j+1} DL + (C+DH)^T \widetilde{R}_{j+1} [A + (1-p)\varepsilon BG]\right]\xi_j \\
&\quad + \xi_j^T \left[L^T D^T \left[I + Q_{j+1}\right] DL + L^T D^T \widetilde{R}_{j+1} [A + (1-p)\varepsilon BG] + A^T \widetilde{S}_{j+1} [A + 2(1-p)\varepsilon BG]\right]\xi_j
\end{aligned}
$$

what finishes the proof. ■

References

[1] Back K., Pliska S.R.: *The Shadow Price of Information in Continuous Time Decision Problems*. "Stochastics" 1987, Vol. 22, pp. 151-186.

[2] Banek T., Kulikowski R.: *Information Pricing in Portfolio Optimization*. "Control and Cybernetics" 2003, Vol. 32, No. 4, pp. 867-882.

[3] Carrillo J.E., Gaimon Ch.: *Improving Manufacturing Performance Through Process Change and Knowledge Creation*. "Management Sciences" 2000, Vol. 46, No. 2, pp. 265-288.

[4] Carr A.N., Hancock P.: *Space and Time in Organizational Change Management*. "Journal of Organizational Change Management" 2006, Vol. 19, No. 5, pp. 545-557.

[5] Carr A.N.: *Jung, Archetypes and Mirroring in Organizational Change Management. Lesson from a longitudinal case study*. "Journal of Organizational Change Management" 2002, Vol. 15, No. 5, pp. 477-489.

[6] Costa L.A., de Matos J.A., e Cunha M.P.: *The Manager as Change Agent*. "International Studies of Management & Organization" 2003-2004, Vol. 33, No. 4, pp. 65-93.

[7] Davis M.H.A., Dempster M.A.H., Elliott R.J.: *On the Value of Information in Controlled Diffusion Processes*. Liber Amicorum for M. Zakai, Academic Press, 1991, pp. 125-138.

[8] Evans L.C., Gariepy R.F.: *Measure Theory and Fine Properties of Functions*. CRC Press, Inc. 1992.

[9] Federer H.: *Geometric Measure Theory*. Springer-Verlag Berlin Heilderberg, 1996.

[10] Ford C.M.: *The Futurity of Decisions as a Facilitator of Organizational Creativity and Change*. "Journal of Organizational Change Management" 2002, Vol. 15, No. 6, pp. 635-646.

[11] Jonnergard K., Karreman M., Svensson C.: *The Impact of Changes in the Corporate Governance System on the Boards of Directors*. "International Studies of Management & Organization" 2004, Vol. 34, No. 2, pp. 113-152.

[12] Lamertz K.: *Comment on "Top executives" Reactions to Change*. "International Studies of Management & Organization" 2004-5, Vol. 34, No. 4, pp. 29-31.

[13] Liptser R.S., Shiryaev A.N.: *Statistics of Random Processes*. Springer-Verlag, Berlin 1977.

[14] Marcus M.L., Robey D.: *Information Technology and Organizational Change: Causal Structure in Theory and Research*. "Management Sciences" 1988, Vol. 34, No. 5, pp. 583-598.

[15] Melkonian T.: *Top Executives' Reactions to Change*. "International Studies of Management & Organization" 2004-2005, Vol. 34, No. 4, pp. 7-28.

[16] Reverdy T.: *Translation Process and Organizational Change*. "International Studies of Management & Organization" 2006, Vol. 36, No. 2, pp. 9-30.

[17] Rishel R.: *A Nonlinear Discrete Time Stochastic Adaptive Control Problem*. Sel. Pap. 7 th. Int. Symp. Math. Theory of Networks Systems, 1985, pp. 585-592.

[18] Schweizer M., Becherer D., Amendinger J.: *A Monetary Value for Initial Information in Portfolio Optimization*. "Finance and Stochastics" 2003, Vol. 7, No. 1, pp. 29-46.

[19] Wets R.J.-B.: *On the Relation between Stochastic and Deterministic Optimization in Control Theory, Numerical Methods and Computer System Modelling*. Lecture Notes in Economics and Mathematical Systems 107. Eds. A. Bensoussen and J.L. Lions. Springer-Verlag, Berlin 1975, pp. 350-361.

[20] Yosida K.: *Fuctional Analysis*. Springer-Verlag, Berlin 1980.

[21] Yu B.T.W., Ming T.W.: *Effects of Control Mechanisms on Positive Organizational Change*. "Journal of Organizational Change Management" 2008, Vol. 21, No. 3, pp. 385-404.

[22] Zabczyk J.: *Chance and Decision, Stochastic Control in Discrete Time*. Scuola Normale Superiore, Quaderni, Pisa 1996.