ISSN 0005-1179, Automation and Remote Control, 2020, Vol. 81, No. 3, pp. 469–482. © Pleiades Publishing, Ltd., 2020.
Russian Text © The Author(s), 2020, published in Avtomatika i Telemekhanika, 2020, No. 3, pp. 114–131.

TOPICAL ISSUE

Simultaneous Impulse and Continuous Control of a Markov Chain in Continuous Time

A. B. Miller[*,**,a], **B. M. Miller**[*,**,***,b], **and K. V. Stepanyan**[*,c]

[*] *Kharkevich Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow, Russia*
[**] *Kazan Federal University, Kazan, Russia*
[***] *Monash University, Melbourne, Victoria, Australia*
e-mail: [a] *amiller@iitp.ru*, [b] *bmiller@iitp.ru*, [c] *KVStepanyan@iitp.ru*

Received June 20, 2019
Revised August 14, 2019
Accepted September 26, 2019

Abstract—We consider continuous and impulse control of a Markov chain (MC) with a finite set of states in continuous time. Continuous control determines the intensity of transitions between MC states, while transition times and their directions are random. Nevertheless, sometimes it is necessary to ensure a transition that leads to an instantaneous change in the state of the MC. Since such transitions require different influences and can produce different effects on the state of the MC, such controls can be interpreted as impulse controls. In this work, we use the martingale representation of a controllable MC and give an optimality condition, which, using the principle of dynamic programming, is reduced to a form of quasi-variational inequality. The solution to this inequality can be obtained in the form of a dynamic programming equation, which for an MC with a finite set of states reduces to a system of ordinary differential equations with one switching line. We prove a sufficient optimality condition and give examples of problems with deterministic and random impulse action.

Keywords: Markov chain, impulse controls, quasi-variational inequality

DOI: 10.1134/S0005117920030066

1. INTRODUCTION

Impulse control as an effect producing instantaneous (in fact, very fast) changes in the state of a dynamical system has been studied in detail since the 1970s. In a series of pioneering works, A. Bensoussan and J.-L. Lions developed the theory of stochastic impulse control, where they generalized the principle of dynamic programming and formulated optimality conditions in the form of a so-called quasi-variational inequality [1]. Their ideas led to the emergence of a new class of discrete-continuous stochastic systems that operate continuously (deterministically or stochastically) between jumps that occur at the moments of application of impulse actions. In stochastic analysis, these ideas were picked up by many researchers who developed the theory of piecewise-deterministic Markov models (PDMM), whose behavior between random possibly controlled moments of jumps is subject to continuous dynamics (see, e.g., [2–4]). Further studies in the field of impulse control led to the emergence of a new class of controlled dynamical systems described by differential equations with measure, which describe in a universal way both impulse and continuous actions. For this class of systems, conditions for the existence of optimal generalized solutions and generalized control measures and simultaneously continuous controls [5, 6] were obtained and optimality conditions were obtained in the form of a generalized maximum principle [7, 8]. Recent years have seen greatly increased interest in the control of Markov chains (MC) and, in particular, in PDMM since these models are better suited for solving optimization problems and are much easier for modeling [9–11]. Moreover, a whole new direction has arisen, namely Markov Decision Processes (MDP), which is very popular in many applied areas, among them being, for instance, the following: control of reservoir systems [12, 13], where impulse effects are manifested in the form of a controlled discharge of water, for example, to avoid overflows; control over the extraction and use of natural resources [14, 15]; control over the Internet to avoid overloads [16, 17]; continuous control of the distribution of water resources by monitoring the prices of water consumed in various sectors of the economy [18, 19]; distribution of gas flows in the gas supply system when it becomes necessary to smooth out seasonal fluctuations in consumption [20]. In all of these areas, impulse actions, as well as continuous ones, arise quite naturally as controls that lead to rapid controlled changes in state; in the control of data transfer this is the release of memory in the event of an overflow risk [21]; in the control of reservoir systems, a controlled discharge of excess water [19, 22].

Another reason for the growing interest in MDP is the relative simplicity and flexibility of modeling continuous and discrete stochastic dynamical systems using MCs with a finite set of states [23]. Since approximation accuracy directly depends on the number of states, the ever-increasing capabilities of available computers allow us to achieve the necessary accuracy in solving optimal control problems without too much complications. In addition, optimal control theory for MCs both in discrete and continuous time is sufficiently developed both for ordinary problems [24] and for problems with state constraints [25–27], so the corresponding numerical procedures are easy to implement. We also note here specifically that for stochastic systems with a continuous set of states, impulse control optimization leads to a special type of optimality condition, namely the quasi-variational inequality [1], which plays the role of a dynamic programming equation. This inequality, even if the existence of its solution is established, is very difficult to solve even numerically since it requires smooth conjugation of solutions for two partial differential equations on an unknown surface. Moreover, the optimal control is a random measure localized on this surface, whose distribution function is singular with respect to the Lebesgue measure, therefore it is possible to determine the characteristics of such a control only with modeling. However, in case of MCs with a finite set of states this problem does not arise. In this case, impulse control is a collection of time pairs and intensity of impulse actions application separated in time, and the problem of solving the corresponding quasi-variational inequality is reduced to solving a system of ordinary differential equations with one switching surface.

This paper is organized as follows. In Section 2, we describe the MC model with continuous and impulse controls; in Section 3 we derive optimality conditions in the form of a quasivariational inequality and provide its solving method. Section 3 gives a proof of the optimality condition sufficiency. Section 4 discusses a numerical example, and Section 5 provides conclusions and directions for further research.

2. MODEL DESCRIPTION AND PROBLEM SETTING

We use the martingale representation of a controlled MC, following the description given in [16, 24, Cpt. 12] for continuous control, supplemented by terms describing impulse effects. All processes are considered on the probability space $\{\Omega, \mathcal{F}, P\}$.

Definition. The process $\{X_t,\ t \in [0,T]\}$ is a controlled Markov process with piecewise constant right-continuous trajectories. The state space of this process is a finite set of unit vectors $\mathbf{S}$ of the form $e_i = (0, 0, \dots, 0, 1, 0, \dots, 0)^{\mathrm{T}} \in \mathbb{R}^n$, i.e., e_i consists of zero elements except for the ith, which is equal to one; the symbol "$^{\mathrm{T}}$" here and below denotes the transposition of vectors and matrices. Thus, $X_t \in \mathbf{S}$.

2.1. Continuous MC Control

Assumption 1. The matrix function $A(t,u)$ with elements $a_{ij}(t,u)$ is a family of MC generators depending on time $t \in [0,T]$, such that the state probability vector $p_t = (p_t^1, \dots, p_t^n)^{\mathrm{T}}$, where $p_t^i = P(X_t = e_i)$, satisfies the direct Kolmogorov equations

$$\frac{dp_t}{dt} = A(t,u)p_t. \tag{2.1}$$

Here the control is $u \in U$, where U is some compact subset of a complete metric space and $A(t,u)$ is continuous on $[0,T] \times U$.

We introduce a right-continuous family of complete σ-algebras generated by the process X_t:

$$\mathcal{F}_t^X = \sigma\{X_s : \ s \in [0,t]\}.$$

Assumption 2. The set $\underline{U}$ of admissible controls $\{u(\cdot)\}$ is the set of $\mathcal{F}_t^X$-predictable processes with values in U. This means that if N_t is the number of state changes for process X, a X_0^t is a sequence of states starting from time point $t = 0$ until the current time $t \in [0,T]$, i.e.,

$$X_0^t = \{(X_0, 0), (X_{\tau_1}, \tau_1), \dots, (X_{\tau_{N_t}}, \tau_{N_t})\}$$

is a collection of states and times of jumps, then for $\tau_{N_t} < t \leqslant \tau_{N_t+1}$ the control $u_t = u(t, X_0^t)$ is a function of X_0^t and current time t [24, 28].

For every control $u(\cdot) \in \underline{U}$ the process $\{X_t\}$ satisfies the stochastic differential equation

$$X_t = X_0 + \int_0^t A(s,u_s)X_s\,ds + W_t, \tag{2.2}$$

where X_0 is the initial condition, and $W_t := \{W_t^1, \dots, W_t^n\}$ is a square integrable $(\mathcal{F}_t^X, P)$ martingale with quadratic characteristic [24, 28, 29]:

$$\left\langle W \right\rangle_t = \int_0^t \mathrm{diag}\ (A(s,u_s)X_s)\ ds + \int_0^t [A(s,u_s)(\mathrm{diag}\ X_s) + (\mathrm{diag}\ X_s)A^{\mathrm{T}}(s,u_s)]\,ds, \tag{2.3}$$

where $\mathrm{diag}(X)$ denotes a diagonal matrix with elements $X^1, \dots, X^n$.

Remark 1. In other words, the process $X(t)$ is a solution to the *martingale problem* (2.2), (2.3) for a controlled Markov chain [24].

2.2. Impulse Control of MC

Impulse control is a set $\mathcal{I} = \{(V_i, \tau_i), i = 1, 2, \dots, N\}$, where $\tau_i < \tau_{i+1} \leqslant T$, $V_i \in \mathcal{V}(X)$ and the number of impulses is at most countable. The use of impulse control at time τ_i leads to an instantaneous random change of state X, so that

$$\Delta X_{\tau_i} = \psi(\tau_i, V_i, X_{\tau_i-}) = \bar{A}(\tau_i, V_i)X_{\tau_i-} + \Delta W_{\tau_i}, \tag{2.4}$$

where

$$\begin{gathered}\mathbf{E}\left[\Delta W_{\tau_i} | \mathcal{F}_{\tau_i-}^X\right] = \mathbf{E}[\Delta W_{\tau_i} | X_{\tau_i-}] = 0, \\ \left\langle \Delta W_{\tau_i} \right\rangle = \mathrm{diag}(\bar{A}(\tau_i, V_i)X_{\tau_i-}) - \bar{A}(\tau_i, V_i)\,\mathrm{diag}(X_{\tau_i-})\bar{A}^{\mathrm{T}}(\tau_i, V_i).\end{gathered} \tag{2.5}$$

Here and below, $\mathbf{E}[\cdot]$ denotes mathematical expectation.

Assumption 3. In (2.4) $\psi(\tau_i, V_i, X_{\tau_i-})$ is a random operator such that for any $X_{\tau_i-} \in \mathbf{S}$, $V_i \in \mathcal{V}(X_{\tau_i-})$ and $\tau_i \in [0, T]$

$$X_{\tau_i} = \Delta X_{\tau_i-} + X_{\tau_i-} \in \mathbf{S}. \tag{2.6}$$

The set $\mathcal{V}(X)$ is compact, the function $\bar{A}(t, V)$ is continuous in t and $\bar{A}(t, 0) = 0$, and therefore $\psi(\tau, 0, X) = 0$ for any $\tau \in [0, T]$ and $X \in \mathbf{S}$.

2.3. Joint Continuous and Impulse Control for an MC

Combining the Eqs. (2.2) and (2.4), we obtain an equation for the joint continuous and impulse control

$$X_t = X_0 + \int_0^t A(s, u(s)) X_s ds + \sum_{\tau_i \leqslant t} \bar{A}(\tau_i, V_i) X_{\tau_i-} + W_t', \tag{2.7}$$

where W_t' is a square integrable martingale with quadratic characteristic

$$\begin{gathered} \langle W_t \rangle = \sum_{\tau_i \leqslant t} \langle \Delta W_{\tau_i} \rangle - \int_0^t (\operatorname{diag} X_{s-}) A^{\mathrm{T}}(s, u(s)) ds \\ - \int_0^t A(s, u(s))(\operatorname{diag} X_{s-}) ds + \operatorname{diag} \int_0^t A(s, u(s)) X_{s-} ds. \end{gathered} \tag{2.8}$$

The control objective is to minimize the performance criterion (2.9)

$$\begin{gathered} J_0[u(\cdot), \mathcal{I}] = \mathbf{E}\left\{ \langle \phi_0, X_T \rangle + \int_0^T \langle g_0(s, u(s)), X_s \rangle ds + \sum_{\tau_i \leqslant T} \langle \psi_0(\tau_i, V_i), X_{\tau_i-} \rangle \right\} \\ = \mathbf{E}\left\{ \sum_{k=1}^N \phi_0^k I\{X_T = e_k\} + \int_0^T g_0^k(s, u(s)) I\{X_s = e_k\} ds + \sum_{\tau_i \leqslant T} \sum_{k=1}^N \psi_0^k(\tau_i, V_i) I\{X_{\tau_i-} = e_k\} \right\} \end{gathered} \tag{2.9}$$

over the set of $\mathcal{F}_t^X$-predictable controls $u(t) \in \underline{U}$ and impulse controls $\mathcal{I}$. In (2.9), $I\{\cdot\}$ denotes the indicator function, i.e., $I\{A\} = 1$ if event A occurs and $I\{A\} = 0$ otherwise.

Assumption 4. The vector function g_0 is continuous in t and u. The vector function $\psi_0(t, V)$ is continuous in t and V and satisfies the inequality $\psi_0^l(t, V) \geqslant C > 0$ for all l and admissible $V \neq 0$. The sets U and $\mathcal{V}$ are compact.

Remark 2. Components of the vector ϕ_0 determine the cost of the terminal state X_T, the vector function g_0 is the integral cost of intermediate states, and ψ_0 is the cost of impulse controls.

Remark 3. In this work, we assume the simplest formulation of the control problem; it can be extended to:

(1) systems with random response to impulse effects;
(2) problems with other types of criteria, such as discounting or averaging over an infinite time interval [21];
(3) problems with state constraints similar to [25–27];
(4) problems with incomplete information [24, 30].

3. DYNAMIC PROGRAMMING EQUATION IN THE FORM OF A QUASIVARIATIONAL INEQUALITY

3.1. Optimality Condition for Joint Continuous and Impulse Control

The cost function in the problem (2.7)–(2.9) has the form

$$\phi(t,X) = \inf_{u(\cdot),\mathcal{I}} \mathbf{E}\left\{ \left\langle \phi_0, X_T \right\rangle + \int\limits_t^T \left\langle g_0(s,u(s)), X_s \right\rangle ds \right.$$
$$\left. + \sum_{t<\tau_i\leqslant T} \left\langle \psi_0(\tau_i, V_i), X_{\tau_i -} \right\rangle \middle| X_{t-} = X \right\} = \left\langle \phi(t), X \right\rangle,$$

where the function $\phi(t) \in \mathbb{R}^N$ is the solution of the quasi-variational inequality

$$\begin{cases} Q_1(t,\phi(t),X) = \left\langle \dfrac{d\phi(t)}{dt}, X \right\rangle + \min\limits_{u\in U} \left\langle A^{\mathrm{T}}(t,u)\phi(t) + g_0(t,u), X \right\rangle \geqslant 0 \\ Q_2(t,\phi(t),X) = \left\langle \phi(t), X \right\rangle - \min\limits_{V\in\mathcal{V}(X)} \left\langle \left[E + \bar{A}^{\mathrm{T}}(t,V)\right]\phi(t) + \psi_0(t,V), X \right\rangle \geqslant 0 \\ \min\left\{Q_1(t,\phi(t),X), Q_2(t,\phi(t),X)\right\} = 0 \end{cases} \tag{3.1}$$

with boundary condition $\phi(T) = \phi_0$. Here and below, E means the identity matrix of the corresponding dimension, and $\mathbf{E}[\cdot]$ still denotes the expectation operator.

Remark 4. Note that relations (3.1) must be satisfied for all values of $X \in \mathcal{S}$, i.e., (3.1) is a system of inequalities that define the value of the cost function for all possible initial conditions t, X.

3.2. Method for Solving the Quasi-Variational Inequality

The quasi-variational inequality (3.1) for a system with a finite set of states allows for a fairly simple method of solution, in contrast to the general case of a system with a continuous part described by the diffusion stochastic equation [1]. The solution, as in the case of a simple dynamic programming equation, is found in inverse time [25–27] with the boundary condition $\phi(T) = \phi_0$ for all components of the vector ϕ_0 satisfying conditions (3.1). Function $\phi(t)$ is determined from the first equation of system (3.1), solving it from the terminal point $\phi(T) = \phi_0$ to the switching line

$$\begin{gathered} 0 = \langle \phi(t), X\rangle - \min_{V\in\mathcal{V}(X)} \left\{\langle \phi(t), [E + \bar{A}(t,V)]X\rangle + \langle \psi_0(t,V), X\rangle\right\} \quad \text{or} \\ G(t,X) = \min_{V\in\mathcal{V}(X)} \langle \bar{A}^{\mathrm{T}}(t,V)\phi(t) + \psi_0(t,V), X\rangle = 0. \end{gathered} \tag{3.2}$$

In impulse control problems, the existence and uniqueness of the solution of the quasi-variational inequality (3.1) play a fundamental role [1]. In the case of control for MCs with a finite set of states, the system of inequalities (3.1) decomposes into a finite number of differential-difference equations corresponding to the number of states. For each value $l = 1, \ldots, N$ we define the switching function $G^l(\phi, t) : \mathbb{R}^N \times [0,T] \to \mathbb{R}^1$ as follows:

$$G^l(\phi,t) = \min_{V\in\mathcal{V}(e_l)} \left\{ \sum_{k=1}^N \phi_k \bar{A}_{l,k}(t,V) + \psi_0^l(t,V) \right\}. \tag{3.3}$$

Due to the continuity of $\bar{A}$ and ψ_0 with respect to t, the switching function is continuous with respect to the pair (ϕ, t) and defines for every l a switching line

$$Sw^l = \left\{(\phi,t,l) : G^l(\phi,t) = 0\right\}. \tag{3.4}$$

Thus, if

$$\phi(t) = \{\phi_1(t), \ldots, \phi_N(t)\}^{\mathrm{T}},$$

then the elements $\phi_l(t), l = 1, \ldots, N$, satisfy the system of equations

$$\frac{d\phi_l(t)}{dt} + \min_{u \in U} \left\{ \sum_{k=1}^{N} \phi_k(t) A_{l,k}(t,u) + g_0^l(t,u) \right\} = 0, \quad \text{if} \quad G^l(\phi(t), t) > 0. \tag{3.5}$$

When solving the first equation of system (3.5) in inverse time with boundary conditions $\phi(T) = \phi_0$ for each $l = 1, \ldots, N$, the switching line is reached at time $\{t^l : G^l(\phi(t^l), t^l) = 0\}$. Due to the condition (3.2), function ϕ remains continuous when crossing the switching line, but the following condition on ϕ allows to determine the value of the impulse control:

$$0 = \min_{V \in \mathcal{V}(\mathcal{X})} \left\{ \left\langle \bar{A}^{\mathrm{T}}(t,V)\phi(t), X \right\rangle + \left\langle \psi_0(t,V), X \right\rangle \right\}. \tag{3.6}$$

The minimizing impulse action in the relation

$$V = \operatorname*{argmin}_{V \in \mathcal{V}(\mathcal{X})} \left\{ \left\langle \bar{A}^{\mathrm{T}}(t,V)\phi(t), X \right\rangle + \left\langle \psi_0(t,V), X \right\rangle \right\} \tag{3.7}$$

exists, and then the solution of system (3.5) continues in reverse time until the next intersection with a switching line. Thus, at the point T_i we use the impulse control V_i defined by relations (3.6), (3.7), and the functions $\phi(t)$ remain continuous. The above procedure determines the solutions of the quasi-variational inequality (3.1).

3.3. A Sufficient Optimality Condition for Joint Impulse and Continuous Control

Proposition 1. *Suppose that:*

a) *functions $\phi(t) \in \mathbb{R}^n$ are determined using the procedure described in Section 3.2 and satisfy the equation*

$$\left\langle \frac{d\phi(t)}{dt}, X \right\rangle + \min_{u \in U(X)} \left\langle A^{\mathrm{T}}(t,u)\phi(t) + g_0(t,u), X \right\rangle = 0 \tag{3.8}$$

on each of the subintervals $[T_i, T_{i+1})$ with terminal condition $\phi(T) = \phi_0$;

b) *controls $u^*(t)$ on $[T_i, T_{i+1})$ are chosen to be Markov, $u^*(t) = u^*(t, X_t)$, and they minimize the right-hand part of the relation (3.8);*

c) *moments of applying the impulse actions $T_i^* = t^*(X)$ are chosen according to the relation*

$$\min_{V \in \mathcal{V}(X)} \left\langle [\bar{A}^{\mathrm{T}}(t^*(X), V)]\phi(t^*(X)) + \psi_0(t^*(X), V), X \right\rangle = 0, \tag{3.9}$$

and impulse action V_i^* are chosen according to relation (3.7).

Then the joint continuous and impulse control $\{u^*(\cdot), T_i^*, V_i^*\}$ is optimal.

Proof. Consider an arbitrary continuous and impulse control $\{u(\cdot), \mathcal{I}\}$ and the corresponding trajectory that satisfies the initial condition $X_t^{u,\mathcal{I}} = X$, and calculate the value

$$\begin{aligned} \Delta J_t^T &= \left\langle \phi(T), X_T^{u,\mathcal{I}} \right\rangle - \left\langle \phi(t), X \right\rangle + \int_t^T \left\langle g_0(s, u(s)), X_s^{u,\mathcal{I}} \right\rangle ds \\ &+ \sum_{t \leqslant T_i \leqslant T} \left\langle \psi_0(T_i, V_i), X_{T_i}^{u,\mathcal{I}} \right\rangle = \left\langle \phi(T), X_T^{u,\mathcal{I}} \right\rangle - \left\langle \phi(t), X \right\rangle + I_1 + I_2, \end{aligned} \tag{3.10}$$

where $\phi(t)$ is the solution of the quasi-variational inequality (3.1), and $X_t^{u,\mathcal{I}}$ is the solution of Eq. (2.7) with initial condition $X_t = X$. Then, using the Ito formula for $X_t^{u,\mathcal{I}}$, we get the relation

$$\Delta J_t^T = \int\limits_t^T \left\langle \frac{d\phi(s)}{ds}, X_s^{u,\mathcal{I}} \right\rangle ds + I_1 + I_2 + \int\limits_t^T \left\langle \phi(s), dX_s^{u,\mathcal{I}} \right\rangle.$$

However,

$$\int\limits_t^T \left\langle \phi(s), dX_s^{u,\mathcal{I}} \right\rangle = \int\limits_t^T \left\langle \phi(s), A(s,u(s)) X_s^{u,\mathcal{I}} \right\rangle ds + \sum_{t \leqslant T_i \leqslant T} \langle \phi(T_i), \bar{A}(T_i, V_i) X_{T_i -} \rangle + \int\limits_t^T \phi(s) dW'(s).$$

Further, for an arbitrary control $u, \mathcal{I}$ according to the relations (3.8), (3.9) we get that

$$\left\langle \frac{d\phi(t)}{dt}, X^{u,\mathcal{I}} \right\rangle + \left\langle A^{\mathrm{T}}(t,u)\phi(t) + g_0(t,u), X^{u,\mathcal{I}} \right\rangle \geqslant 0, \tag{3.11}$$

$$\left\langle \phi(T_i), \bar{A}(T_i, V_i) X_{T_i -} + \psi_0(T_i, V_i), X_{T_i -} \right\rangle \geqslant 0. \tag{3.12}$$

Since the integral over the martingale has zero expectation, combining (3.11) and (3.12) and substituting into (3.10), we obtain for an arbitrary control the inequality

$$\mathbf{E}\left\{ \left\langle \phi(T), X_T^{u,\mathcal{I}} \right\rangle + \int\limits_t^T \left\langle g_0(s), u(s)), X_s^{u,\mathcal{I}} \right\rangle ds + \sum_{t \leqslant T_i \leqslant T} \left\langle \psi_0(T_i, V_i), X_{T_i}^{u,\mathcal{I}} \right\rangle \right\} \geqslant \left\langle \phi(t), X \right\rangle. \tag{3.13}$$

But if we use a control satisfying the optimality conditions, equality is achieved in the inequality (3.13). This completes the proof of the proposition.

3.4. Problems with Constraints

Problems with constraints undoubtedly deserve a separate publication; here we show only a sketch of a possible approach. Suppose that a set of criteria (3.14) is given, and a feasible solution must satisfy the constraints

$$J_k[u(\cdot), \mathcal{I}] \leqslant 0 \quad \text{for all} \quad k = 1, \ldots, M,$$

where

$$J_k[u(\cdot), \mathcal{I}] = \mathbf{E}\left\{ \left\langle \phi_k, X_T \right\rangle + \int\limits_0^T \left\langle g_k(s.u(s)), X_s \right\rangle ds + \sum_{\tau_i \leqslant T} \left\langle \psi_k((\tau_i, V_i), X_{\tau_i -} \right\rangle \right\}. \tag{3.14}$$

Assuming that the set of admissible trajectories is nonempty, and assuming that the set of controls is extended to the set of control measures, we can use the Lagrangian minimization procedure, which is applicable if the set of admissible values of the criteria $\{J_k\} \in \mathbb{R}^M$ is convex. This property, as shown in [25–27], holds for control problems with continuous controls; however, for the problem of joint continuous and impulse control it can also be established for an extended class of control measures. A rather lengthy proof mainly repeats the arguments given in [25, 26], and will be the subject of a separate publication.

4. EXAMPLES WITH IMPULSE CONTROLS

4.1. Deterministic Continuous Control

In this work we consider the model of a service system with an extension that allows impulse control, with $N = 4$ states:

- 1—“load is below normal,”
- 2—“normal,”
- 3—“critically high, immediate buffer cleaning required,”
- 4—“overflow, leading to a service outage or requiring immediate cleaning of the claim queue and reducing the load to normal.”

Admissible transitions corresponding to continuous dynamics:

- $1 \to 2, 2 \to 3, 3 \to 4$, which occur with intensity $\lambda(t)$ of the input stream of claims;
- $2 \to 1, 3 \to 2, 4 \to 3$, which occur with the intensity $\mu(t) \in [0, U(X_t)]$ of service claims;
- transitions corresponding to impulse control are admissible in states 3, 4 and transfer the system from state $3 \to 2$ and $4 \to 2$;
- random transitions are possible, for example, when using impulse control from state $4 \to 2$ or $4 \to 3$ with probabilities p_1 and $1 - p_1$ respectively.

State 4 is critical, and the time spent in it should be minimized. The control objective is to “maintain” the system in state 2, while keeping the time spent in states $3, 4$ at an acceptable level. The state vector $X \in \mathbb{R}^4$ and the matrix of transition intensities $A(t, \mu)$, which depends on the “continuous” control μ, has the form (for simplicity we omit the dependencies on time t)

$$A(t,\mu) = \begin{pmatrix} -\lambda & \mu_2 & 0 & 0 \\ \lambda & -(\lambda+\mu_2) & \mu_3 & 0 \\ 0 & \lambda & -(\lambda+\mu_3) & \mu_4 \\ 0 & 0 & \lambda & -\mu_4 \end{pmatrix}.$$

Variables μ_2, μ_3, μ_4 correspond to the intensity of claim processing in states 2–4 respectively.

The model proposed in this work is similar both to the problems of managing the water level in a reservoir [19, 31, 32] and problems of managing data flows in the Internet [16, 17]. This analogy is not so artificial, since in both cases the increase in the level (of water or load) occurs randomly due to rains or the input flow of claims, and decrease in the level occurs due to evaporation and/or consumption by various consumers: industry, agriculture or settlements, if we are talking about water, or by completing tasks if we are talking about a service system. In the former case, impulse control corresponds to a controlled discharge to avoid the overflow of the reservoir; in the latter case, to flushing the task buffer in order to avoid the congestion. In addition to impulse control, “continuous” control is also used, which reduces the level in the former case due to the supply of consumers and in the latter, by reducing the number of claims awaiting processing. Both models have certain features, for example, service intensity may depend on the length of the queue, but in the case of reservoirs the intensity of the reduction due to evaporation may depend on seasonal changes and the current level.

4.2. Deterministic Impulse Control

We assume that impulse control can be used in states 3 and 4 with transitions $3 \to 2$ and $4 \to 2$.

The transition matrix corresponding to the application of control V_3, in state 3 at some time moment τ_i is equal to

$$E + \bar{A}(\tau_i, V_3) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

and the matrix corresponding to the application of control V_4, in state 4 at some time moment τ_i is respectively

$$E + \bar{A}(\tau_i, V_4) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

If the impulse control cost is chosen as

$$\psi_0(t, V) = (0, 0, 1, 1)^{\mathrm{T}}, \tag{4.1}$$

then the condition for applying impulse control at the point T_i as a function of $\phi(T_i)$ according to (3.2) is given by

$$0 \leqslant \langle \phi(T_i), \bar{A}(T_i, V_i)X \rangle + \langle \psi_0(T_i, V_i), X \rangle,$$

which yields relations for the values ϕ_2, ϕ_3, ϕ_4

$$\begin{aligned} \phi_3(T_i) &\geqslant \phi_2(T_i) + 1, \text{ if } X = e_3, \\ \phi_4(T_i) &\geqslant \phi_2(T_i) + 1, \text{ if } X = e_4. \end{aligned} \tag{4.2}$$

The current cost characterizing a deviation from the "normal" state 2 and service cost is

$$g_0(t, \mu) = \begin{pmatrix} 1 \\ 0 + \mu_2^2 \\ 1 + \mu_3^2 \\ 2 + \mu_4^2 \end{pmatrix}. \tag{4.3}$$

Here $\mu_i \in [0, U(e_i)]$ for $i = 2, 3, 4$, which corresponds to constraints on the controls, i.e., service intensities in the corresponding states.

Thus, the continuous part of the system of Eqs. (3.1) has the form

$$\begin{cases} \dfrac{d\phi_1}{dt} = -\lambda\phi_1 + \mu_2\phi_2 + 1 \\ \dfrac{d\phi_2}{dt} = \min\limits_{\mu_2 \in [0, U(e_2)]} [\lambda\phi_1 - (\lambda + \mu_2)\phi_2 + \mu_3\phi_3 + \mu_2^2] \\ \dfrac{d\phi_3}{dt} = \min\limits_{\mu_3 \in [0, U(e_3)]} [\lambda\phi_2 - (\lambda + \mu_3)\phi_3 + \mu_4\phi_4 + \mu_3^2 + 1] \\ \dfrac{d\phi_4}{dt} = \min\limits_{\mu_4 \in [0, U(e_4)]} [\lambda\phi_3 - \mu_4\phi_4 + 2 + \mu_4^2]. \end{cases} \tag{4.4}$$

System (4.4) is solved from the terminal point T with the boundary condition (4.1) and, even if the condition for applying the impulse control is satisfied, its solution remains continuous. Further, when implementing the solution, if the system is in state 3 or 4 and condition (4.2) holds, then impulse control is applied and the system goes into state 2. Since in state 2 impulse control is inadmissible, continuous control is applied until the system goes into state 3 or 4 and the condition for applying impulse control is met. State 1 is not unreachable. The system can appear in it both because of the initial conditions and due to the "natural" dynamics.

4.3. Stochastic Impulse Control

In this case, after applying impulse control, transitions from state 4 to states 2 and 3 are possible with probabilities p_1 and $1-p_1$ respectively. We assume that the rest of the parameters do not change. Then the matrix corresponding to the application of impulse control has the form

$$E+\bar{A}(\tau_i,V_i)=\begin{pmatrix}1&0&0&0\\0&1&0&p_1\\0&0&1&1-p_1\\0&0&0&0\end{pmatrix}.$$

The condition for the possibility of using impulse action is

$$\phi_4(T_i)\geqslant p_1\phi_2(T_i)+(1-p_1)\phi_3(T_i)+1.$$

Remark 5. We note the remarkable fact that the function $\phi(t)$ remains continuous at the points where impulse control is applied. This, however, is not all that surprising since, as shown in [25–27], components of this function coincide with the conjugate variables in the maximum principle, and therefore they are independent of the control, as in the case of the usual impulse control problem [6–8].

A similar example was considered in [33].

4.4. Implementation of Joint Continuous and Impulse Control

We have to emphasize the difference between systems with continuous and discrete sets of states. In the first case, impulse control is usually applied at the initial moment of time if the system is located above the switching surface. If then, as a result of random evolution, the system again comes to the discontinuity surface, then the impulse control acts to keep the system below the switching surface, implementing the so-called singular repulsion control [34]. In this case, the impulses are not separated and a singular control of the Cantor ladder type arises. Controls of this type are characteristic of impulse control problems for stochastic diffusion type processes. In the case of Markov chain control, if the condition

$$\mathcal{V}(X_{\tau_i})=\emptyset$$

holds, then after a jump instantaneous application of impulse control is impossible, and the next impulse control is possible only when the system again falls into a state where

$$\mathcal{V}(X_{t-})\neq\emptyset.$$

However, if as a result of impulse control having a random or deterministic character, the system either remains on the switching surface or falls into a state where impulse control is again possible, then the impulses can be repeated until the system enters a state where impulse action is impossible. Thus, the effect of a “multiple” impulse (burst of impulses) arises [21]. In the considered example, upon transition from state 4 to 3, this effect is possible if in state 3 the impulse action is again admissible. In this case, the transition from state 4 to 2 occurs in two jumps $4\to 3\to 2$, but the cost of such an impulse effect is equal to the total cost of two transitions, i.e., $1+1=2$.

4.5. Examples of Finding “Continuous” and Impulse Controls

As an illustrative example, consider the deterministic impulse control system described in Section 2.3 with the following parameters:

- $T=1$;

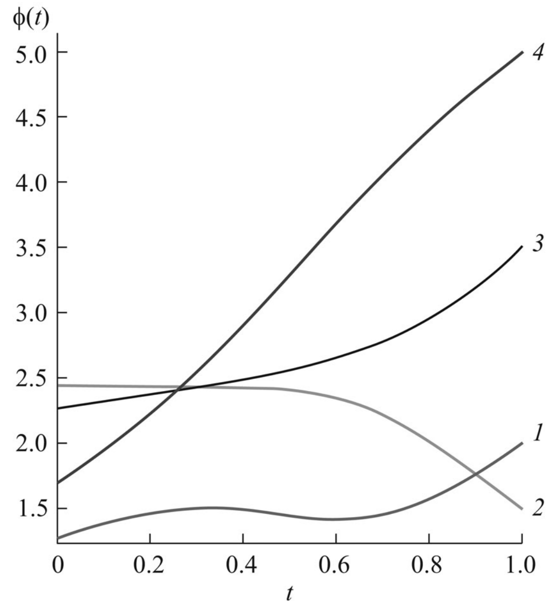

Fig. 1. Solution of the system (4.6) in reverse time with terminal conditions (4.5). *1*—$\phi_1(t)$, *2*— $\phi_2(t)$, *3*—$\phi_3(t)$,*4*—$\phi_4(t)$.

- terminal conditions corresponding to the "penalty" for staying in states 1–4 at a finite time moment are equal to

$$\phi_0 = \phi(T) = (2.0; 1.5; 3.5; 5.0)^{\mathrm{T}}; \tag{4.5}$$

- constraints on the control or service rate $\mu_2^{\max} = U(e_2) = 0.2$, $\mu_3^{\max} = U(e_3) = 0.4$ are imposed only for states 2 and 3, and state 4 corresponds to stopping the service, therefore state 4 is assumed to be absorbing.

Thus, the continuous part of the system of Eqs. (3.1) in this case has the form

$$\begin{cases} \dfrac{d\phi_1}{dt} = -\lambda\phi_1 + \mu_2\phi_2 + 1 \\ \dfrac{d\phi_2}{dt} = \min\limits_{\mu_2 \in [0, \mu_2^{\max}]} [\lambda\phi_1 - (\lambda + \mu_2)\phi_2 + \mu_3\phi_3 + \mu_2^2] \\ \dfrac{d\phi_3}{dt} = \min\limits_{\mu_3 \in [0, \mu_3^{\max}]} [\lambda\phi_2 - (\lambda + \mu_3)\phi_3 + \mu_4\phi_4 + \mu_3^2 + 1] \\ \dfrac{d\phi_4}{dt} = \lambda\phi_3 + 2. \end{cases} \tag{4.6}$$

Solution of system (4.6) in inverse time for the intensity of the input stream of claims defined by the relation

$$\lambda(t) = 0.25\cos(2\pi t) + 0.5, \tag{4.7}$$

is shown in Fig. 1.

The switching line (3.2) for an admissible impulse control is calculated as $SW(t) = \phi_3(t) - \phi_2(t) - 1 = 0$ and is shown in Fig. 2. Thus, if the system is in state 3 and condition $SW(t) \geqslant 0$ is fulfilled, then impulse control should be used, which will transfer the system to state 2. Thus, the domain of impulse control is simultaneously state 3 and time interval $\approx [0.8; 1.0]$.

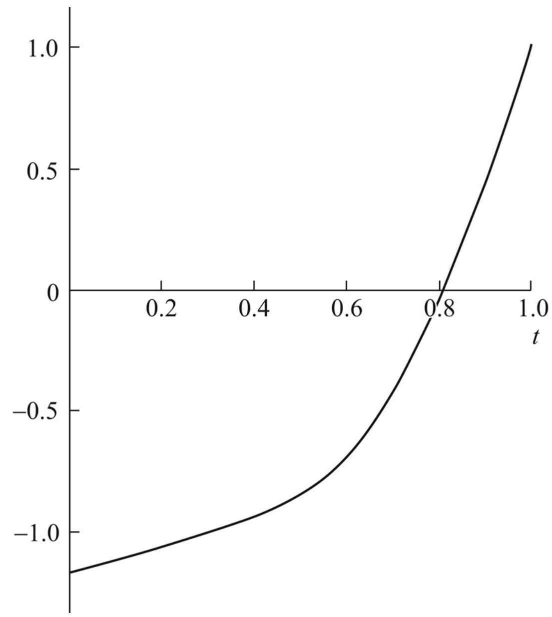

Fig. 2. Switching line for impulse control in state 3. The area where $SW(t) \geqslant 0$ is the domain of impulse control.

Note that the use of impulse control allows to reduce the value of the quality criterion, although this depends on the probability of getting into state 3 on the interval $[0.8; 1.0]$. If this happens, for example, at time point $t = 0.8$, then the value of the quality criterion for state 3 at the end time of the process $t = 1.0$ is $\phi_2(1.0) + 1 \approx 2.5$, which is less than without applying impulse control, i.e., $\phi_3(1.0) \approx 3.5$. However, in order to calculate the value of the performance criterion when using both continuous and impulse controls it is necessary to calculate the value of the problem cost taking into account the random nature of the impulse control application, which should be the subject of another publication. Another important issue requiring separate consideration is the calculation of the state distribution function under impulse control; this leads to the need to solve direct Kolmogorov equations, which in this case are no longer ordinary differential equations but rather discrete-continuous equations with random switching times.

5. CONCLUSION

In this work, we have developed an approach to the numerical solution of the stochastic control problem for MCs with "continuous" and impulse controls. An important feature of the MC model is a simpler procedure for determining optimality conditions than for stochastic systems with a continuous set of states. Further research will be aimed at solving the applied problems of natural resource management using MC models, for example, to control the water supply [19], the distribution of natural gas [20], control the transmission of data through telecommunication networks with unstable characteristics, and, in particular, for communication and data transfer with unmanned aerial vehicles [30, 35].

FUNDING

The work done by A.B. Miller and B.M. Miller was partially financially supported within the framework of State support for the Kazan (Volga) Federal University in order to increase its competitiveness among the world's leading research and educational centers.

REFERENCES

1. Bensoussan, A. and Lions, J.L., *Contrôle Impulsionnel et Inéquations Quasivariationnelles*, Paris: Dunod, 1982. Translated under the title *Impul'snoe upravlenie i kvazivariatsionnye neravenstva*, Moscow: Nauka, 1987.
2. Yushkevich, A.A., Markov Decision Processes with Both Continuous and Impulsive Control, in *Stochastic Optimization. Lect. Notes in Control and Information Sci.*, Arkin, V.I., Shiraev, A., and Wets, R., Eds., Berlin: Springer, 1986, vol. 81, pp. 234–246. https://doi.org/10.1007/BFb0007100
3. Gatarek, D., Impulsive Control of Piecewise-Deterministic Processes, in *Stochastic Systems and Optimization. Lect. Notes in Control and Information Sci.*, Zabczyk, J., Ed., Berlin: Springer, 1989, vol. 136, pp. 298–308. https://doi.org/10.1007/BFb0002690
4. Dempster, M.A.H. and Ye, J.J., Impulse Control of Piecewise Deterministic Markov Processes, *Ann. Appl. Prob.*, 1995, vol. 5, no. 2, pp. 399–423. https://doi.org/10.1214/aoap/1177004771
5. Dufour, F. and Miller, B.M., Generalized Solutions in Nonlinear Stochastic Control Problems, *SIAM J. Control Optim.*, 2002, vol. 40, pp. 1724–1745. https://doi.org/10.1137/S0363012900374221
6. Miller, B.M. and Rubinovich, E.Ya., *Impulsive Control in Continuous and Discrete-Continuous Systems*, New York: Kluwer, 2003. https://doi.org/10.1007/978-1-4615-0095-7
7. Miller, B.M. and Rubinovich, E.Ya., *Optimizatsiya dinamicheskikh sistem s impul'snymi upravleniyami* (Optimization of Dynamical Systems with Impulse Controls), Moscow: Nauka, 2005.
8. Miller, B.M. and Rubinovich, E.Ya., *Optimizatsiya dinamicheskikh sistem s impul'snymi upravleniyami i udarnymi vozdeistviyami* (Optimization of Dynamical Systems with Impulse Controls and Impact Influences), Moscow: LENAND/URSS, 2019.
9. Dufour, F., Horiguchi, M., and Piunovskiy, A.B., Optimal Impulsive Control of Piecewise Deterministic Markov Processes, *Stoch.*, 2016, vol. 88, no. 7, pp. 1073–109. https://doi.org/10.1080/17442508.2016.1197925
10. Shujun Wang and Zhen Wu, Maximum Principle for Optimal Control Problems of Forward-Backward Regime-Switching Systems Involving Impulse Controls, *Math. Probl. Eng.*, 2015, Article ID 892304. http://dx.doi.org/10.1155/2015/892304
11. Dufour, F. and Piunovskiy, A.B., Impulsive Control for Continuous-Time Markov Decision Processes: A Linear Programming Approach, *Appl. Math. Optim.*, 2016, vol. 74, pp. 129–161. https://doi.org/10.1007/s00245-015-9310-8
12. Delebecque, F. and Quadrat, J.-P., Optimal Control of Markov Chains Admitting Strong and Weak Interactions, *Automatica*, 1981, vol. 17, no. 2, pp. 281–296. https://doi.org/10.1016/0005-1098(81)90047-9
13. Bae, J., Kim, S., and Lee, E.Y., Average Cost under the $p^M_{\lambda,\tau}$ Policy in a Finite Dam with Compound Poisson Inputs, *J. Appl. Probab.*, 2003, vol. 40, no. 2, pp. 519–526. https://doi.org/10.1239/jap/1053003561
14. Filar, J.A. and Vrieze, K., *Competitive Markov Decision Processes—Theory, Algorithms and Applications*, New York: Springer, 1997. https://doi.org/10.1007/978-1-4612-4054-9
15. Williams, B.K., Markov Decision Processes in Natural Resources Management: Observability and Uncertainty, *Ecological Modelling*, 2009, vol. 220, no. 6, pp. 830–840. https://doi.org/10.1016/j.ecolmodel.2008.12.023
16. Miller, A., Dynamic Access Control with Active Users, *J. Commun. Technol. Electron.*, 2010, vol. 55, no. 12, pp. 1432–1441. https://doi.org/10.1134/S1064226910120168
17. Miller, A. and Miller, B., Control of Connected Markov Chains. Application to Congestion Avoidance in the Internet, *Proc. 50th IEEE Conf. on Decision and Control and European Control Conf. (CDC-ECC)*, Orlando, USA, 2011, pp. 7242–7248. https://doi.org/10.1109/CDC.2011.6161029
18. Miller, B. and McInnes, D., Optimal Management of a Two Dam System via Stochastic Control: Parallel Computing Approach, *Proc. 50th IEEE Conf. on Decision and Control and Eur. Control Conf. (CDC-ECC)*, Orlando, USA, December 12–15, 2011, pp. 1417–1423. https://doi.org/10.1109/CDC.2011.6160566

19. McInnes, D. and Miller, B., Optimal Control of Large Dam Using Time-Inhomogeneous Markov Chains with an Application to Flood Control, *IFAC-PapersOnLine*, 2017, vol. 50, no. 1, pp. 3499–3504. https://doi.org/10.1016/j.ifacol.2017.08.936

20. McInnes, D., Miller, B., and Schreider, S., Optimisation of Gas Flows in South Eastern Australia via Controllable Markov Chains, *Proc. Australian Control Conf.*, Newcastle: Engineers Australia, 2016, November 3–4, pp. 341–346. https://doi.org/10.1109/AUCC.2016.7868213

21. Avrachenkov, K., Habachi, O., Piunovskiy, A., and Zhang, Y., Infinite Horizon Optimal Impulsive Control with Applications to Internet Congestion Control, *Int. J. Control*, 2015, vol. 88, no. 4, pp. 703–716. https://doi.org/10.1080/00207179.2014.971436

22. McInnes, D. and Miller, B., Optimal Control of Time-Inhomogeneous Markov Chains with Application to Dam Management, *Proc. Australian Control Conf.*, Perth, Australia, 2013, pp. 230–237. https://doi.org/10.1109/AUCC.2013.6697278

23. Kushner, H.J. and Dupuis, P.G., *Numerical Methods for Stochastic Control Problems in Continuous Time*, New York: Springer Science+Business Media, 2001, 2nd ed. https://doi.org/10.1007/978-1-4613-0007-6

24. Elliott, R.J., Aggoun, L., and Moore, J.B., *Hidden Markov Models. Estimation and Control*, New York: Springer-Verlag, 2008. https://doi.org/10.1007/978-0-387-84854-9

25. Miller, B., Miller, G., and Semenikhin, K., Torwards the Optimal Control of Markov Chains with Constraints, *Automatica*, 2010, vol. 46, pp. 1495–1502. https://doi.org/10.1016/j.automatica.2010.06.003

26. Miller, B.M., Miller, G.B., and Semenikhin, K.V., Methods to Design Optimal Control of Markov Process with Finite State Set in the Presence of Constraints, *Autom. Remote Control*, 2011, vol. 72, no. 2, pp. 323–341.

27. Miller, B.M., Miller, G.B., and Semenikhin, K.V., Optimal Control Problem Regularization for the Markov Process with Finite Number of States and Constraints, *Autom. Remote Control*, 2016, vol. 77, no. 9, pp. 1589–1611. https://doi.org/10.1134/S0005117916090071

28. Brémaud, P., *Point Processes and Queues: Martingale Dynamics*, New York: Springer, 1981.

29. Liptser, R.Sh. and Shiryaev, A.N., *Statistics of Random Processes*, New York: Springer, 1979.

30. Miller, B., Miller, G., and Semenikhin, K., Optimization of the Data Transmission Flow from Moving Object to Nonhomogeneous Network of Base Stations, *IFAC-PapersOnLine*, 2017, vol. 50, no. 1, pp. 6160–6165. https://doi.org/10.1016/j.ifacol.2017.08.981

31. Miller, A., Miller, B., Popov, A., and Stepanyan, K., Towards the Development of Numerical Procedure for Control of Connected Markov Chains, *Proc. 5th Australian Control Conf. (AUCC)*, Gold Coast, Australia, November 5–6, 2015, 2015, pp. 336–341.

32. Miller, B., Stepanyan, K., Miller, A., and Popov, A., Towards One Nonconvex Connected Markov Chain Control Problem. An Approach to Numerical Solution, *2018 Australian New Zealand Control Conf. (ANZCC)*, Swinburne University of Technology, Melbourne, Australia, Dec. 7–8, 2018, pp. 172–177. https://doi.org/10.1109/ANZCC.2018.8606576

33. Miller, A., Miller, B., and Stepanyan, K., A Numerical Approach to Joint Continuous and Impulsive Control of Markov Chains, *IFAC-PapersOnLine*, 2018, vol. 51, no. 32, pp. 462–467. https://doi.org/10.1016/j.ifacol. 2018.11.428

34. Karatzas, I. and Shreve, S.E., Connection Between Optimal Stopping and Singular Stochastic Control I. Monotone Follower Problems, *SIAM J. Control Optim.*, 1984, vol. 22, no. 6, pp. 856–877. https://doi.org/10.1137/0322054

35. Miller, B.M., Miller, G.B., and Semenikhin, K.V., Optimal Channel Choice for Lossy Data Flow Transmission, *Autom. Remote Control*, 2018, vol. 79, no. 1, pp. 66–77. https://doi.org/10.1134/S000511791801006X

This paper was recommended for publication by E.Ya. Rubinovich, a member of the Editorial Board