Article

Towards Tensor Representation of Controlled Coupled Markov Chains

Daniel McInnes [1] , Boris Miller [1,2] , Gregory Miller [3,]* and Sergei Schreider [4]

[1] School of Mathematics, Monash University, Wellington Road, Clayton VIC 3800, Australia;
    daniel.mcinnes@monash.edu (D.M.); bmiller@iitp.ru (B.M.)
[2] Institute for Information Transmission Problems RAS, 19/1 Bolshoy Karetny Per., 127051 Moscow, Russia
[3] Institute of Informatics Problems of Federal Research Center "Computer Science and Control" RAS,
    44/2 Vavilova Str., 119333 Moscow, Russia
[4] Department of Management and Information Systems, Rutgers Business School, Rutgers—The State
    University of New Jersey, New Brunswick, NJ 07102, USA; Sergei.Schreider@rutgers.edu
* Correspondence: gmiller@frccsc.ru

check for
updates

Abstract: For a controlled system of coupled Markov chains, which share common control parameters, a tensor description is proposed. A control optimality condition in the form of a dynamic programming equation is derived in tensor form. This condition can be reduced to a system of coupled ordinary differential equations and admits an effective numerical solution. As an application example, the problem of the optimal control for a system of water reservoirs with phase and balance constraints is considered.

Keywords: coupled markov chains; stochastic control; optimal control; tensor representation; dynamic programming

1. Introduction

The analysis and optimization of systems of coupled Markov Chains (MC) appear in various applied areas such as environmental management [1,2], control of dams [3–5], control of data transmission systems for the avoidance of congestion [6–8], and many others. In general, the optimal control of MC with constraints and various criteria leads to the solution of a system of ordinary differential equations. However, the coupled MC produces serious difficulties due to a dramatic increase in the number of states, which renders impossible the numerical solution of the optimization problem with the aid of standard computers in a reasonable time [9]. Moreover, the presence of constraints very often leads to rather complicated and cumbersome mathematical problems.

The evolution of single controlled MC may be effectively represented via a vector-valued function satisfying stochastic differential equations with a control dependent generator [10]. If a set of MCs shares common control parameters, they no longer can be analyzed separately, since the system becomes coupled. Approaching the coupled system as a whole leads to the necessity of the consideration of a global system state, which incorporates the states of all individual MCs involved. Joining together the states of all MCs as a set of vectors (perhaps of different dimensions) leads to the description of their evolution in tensor form. The next step of the control synthesis for this coupled system is the derivation of the optimality conditions in the form of a special dynamic programming equation (DPE). The demand for accuracy leads, however, to the extension of MCs' state space, and by that, to the increase of the dimensionality of the DPE. At the same time, the numerical solution of DPE allows the parallelization of the most time-consuming operations, and this is one of the advantages of the tensor representation. Here we give an algorithm for DPE derivation, which is the principal

step for its numerical solution. It should be noted that this paper follows the previously published works [6,9,11] and provides a proper mathematical justification for the results presented there.

As an example, we consider the management of coupled dams under non-stationary seasonally changing random inflows/outflows. The presentation of a dam's state as a continuous-time MC permits one to take into account the random character of the incoming and outgoing water flow due to rain, evaporation, and, of course, customer demands. However, sometimes it is necessary to organize the interflow between the different parts of a system, such as controlled flow from one dam to another or controlled release, to avoid overflow. In all these cases, it is necessary to consider the system as a whole, that is, the global state of the dam system. The current water level of each dam is described by the state of a continuous-time MC, hence the state of the whole system of the coupled dams is represented in tensor form. The connection of MCs is a result of the controlled flow between the dams. The control aims to maintain the required water levels in given weather conditions, and at the same time, to satisfy the customer demands. The general approach is based on the solution of DPE in tensor form. This equation may be reduced to a system of ordinary differential equations. We suggest here the procedure for the generation of this system and also the approach to the minimization in its right-hand side (RHS), which may be realized for each state of the coupled MC independently.

The structure of the paper is as follows: in Section 2 we describe the model of the controlled MC system. The martingale representation and stochastic evolution of the global state of coupled MC is given in Section 3. The optimal control and Kolmogorov equation for the distribution of states of the controlled coupled system of MC is in Section 4. A numerical example motivated by the Goulburn River water cascade is presented in Section 5.

2. Model of Controlled Markov Chains System

In this section we give the definition of the model of the controlled coupled MCs. Let us consider a set of M MCs. The m-th MC has N^m possible states and a time- and control-dependent generator $A^m(t, u) = [a_{i,j}^m(t, u)]$, $i, j = 1, \ldots, N^m$:

- $a_{i,j}^m(t, u) \geq 0$, for $i \neq j$;
- $a_{i,i}^m(t, u) = - \sum\limits_{j \neq i} a_{j,i}^m(t, u)$.

The controlparameter $u \in U$, where U is a compact set in some complete metric space, and the matrix valued functions A^m, $m = 1, \ldots, M$ are assumed to be continuous in both parameters $(t, u) \in [0, T] \times U$, where $T < \infty$. The control parameter u is shared by all the MCs, and hence defines their interconnection.

2.1. Global State

Let the local state space of the m-th MC be given by a set of standard basis vectors $X_t^m \in \mathcal{S}^m = \{e^1, \ldots, e^{N^m}\}$, then its dynamics can be described by the following stochastic differential equation [10]:

$$X_t^m = X_0^m + \int\limits_0^t A^m(s, u(s)) X_s^m ds + W_t^m, \quad m = 1, \ldots, M, \tag{1}$$

where X_0^m is the initial state of the m-th MC and the process W_t^m is a square integrable martingale with bounded quadratic variation. We assume that the martingales W_t^l and W_t^m are independent for $m \neq l$, and hence the MCs do not experience simultaneous transitions.

The first approach to the controlled coupled MCs was presented in References [6,9,11], where the global system state was described as a tensor product of the individual MCs' states:

$$\mathbf{X}_t = X_t^1 \otimes X_t^2 \otimes \ldots \otimes X_t^M = \bigotimes_{m=1}^M X_t^m, \tag{2}$$

$\mathbf{X}_t \in \mathbf{S}$, $\mathbf{S} = \mathcal{S}^1 \times \ldots \times \mathcal{S}^M$, where $\times$ denotes the Cartesian product (the definition of tensor product can be found in Appendix A). For the combination of local states $X_t^1 = e^{i^1}, \ldots, X_t^M = e^{i^M}$, the global state $\mathbf{X}_t = \mathbf{e}^{i^1 \ldots i^M} = e^{i^1} \otimes \ldots \otimes e^{i^M}$ can be represented as a multidimensional array of shape $(N^1, \ldots, N^M)$ with single non-zero element $\mathbf{X}_t^{i^1 \ldots i^M} = 1$. This representation allows us to significantly simplify the expressions involved in the DPE and makes the implementation of the control optimization algorithms rather straightforward, especially with programming tools that allow multidimensional array objects (e.g., Matlab, Python NumPy or Julia).

Let the set of admissible controls $\mathcal{U}$ be defined as a set of $\mathcal{F}_t^{\mathbf{X}}$-predictable controls taking values in U (here $\mathcal{F}_t^{\mathbf{X}}$ stands for a natural filtration associated with the stochastic process $\mathbf{X}_t$). This assumption ensures the measurability of the admissible controls. In other words, if the number of global state transitions of the coupled MC system up to the current time $t \in [0, T]$ is $\mathcal{N}_t$ and τ_k is the time of the k-th jump and

$$X^m|_0^t = \{(X_0^m, 0), (X_1^m, \tau_1), \ldots, (X_{\mathcal{N}_t}^m, \tau_{\mathcal{N}_t})\}$$

is the history of the m-th MC on the time interval $[0, t]$ (set of state and jump time pairs), then for $\tau_{\mathcal{N}_t} \le t < \tau_{\mathcal{N}_{t+1}}$ the controls $u(t) = u(t, \mathbf{X}|_0^t)$ are measurable with respect to t and the global history $\mathbf{X}|_0^t = \{X^0|_0^t, \ldots, X^M|_0^t\}$ [9].

2.2. Performance Criterion

Let $f(s, u(s), \mathbf{X}_s)$ be a running cost function defined for the global state $\mathbf{X}_s$ and time $s \in [0, T]$, and $\varphi_0(\mathbf{X}_T)$ be a terminal condition. Then a general performance criterion to be minimized is given by

$$J[u(\cdot), \mathbf{X}(\cdot)] = \mathbb{E}\left[\varphi_0(\mathbf{X}_T) + \int_0^T f(s, u(s), \mathbf{X}_s)ds\right]. \tag{3}$$

Here $f(s, u(s), \cdot)$ is, in fact, a multidimensional array-valued function defined for any possible global state of the coupled MC system $\mathbf{X}_t = \mathbf{e}^{i^1 \ldots i^M} = e^{i^1} \otimes \ldots \otimes e^{i^M}$. For example, the running cost function can be represented as

$$f(s, u(s), \mathbf{X}_t) = \langle \mathbf{f}(s, u(s)), \mathbf{X}_t \rangle = \sum_{i^1 \ldots i^M} \mathbf{f}_{i^1 \ldots i^M}(s, u(s)) \mathbf{X}_t^{i^1 \ldots i^M} = \sum_{i^1 \ldots i^M} \mathbf{f}_{i^1 \ldots i^M}(s, u(s)) I\{\mathbf{X}_t^{i^1 \ldots i^M} = 1\}, \tag{4}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product and $I\{\cdot\}$ is the indicator function. Thus, for given t and $u(t)$, the state $\mathbf{X}_t = \mathbf{e}^{i^1 \ldots i^M}$ selects a single value $\mathbf{f}_{i^1 \ldots i^M}(s, u(s))$ from the multidimensional array $\mathbf{f}(s, u(s))$. The same is valid for the terminal condition $\varphi_0(\mathbf{X}_T) = \langle \boldsymbol{\varphi}_T, \mathbf{X}_T \rangle$, which is defined by a multidimensional array $\boldsymbol{\varphi}_T$.

Assumption 1. *The elements of $\mathbf{f}(s, u(s))$ are bounded from below and are continuous functions on $[0, T] \times U$.*

3. Martingale Representation of the Global State

In this section we derive the martingale representation for the global state of the controlled coupled MC system analogous to (1). To that end, we need to introduce some additional notations. Let I^{A^m} denote an identity matrix of the same shape as the matrix $A^m(s, u(s))$, $m = 1, \ldots, M$, and $A \otimes B$ denote the tensor product of matrices A and B (see Appendix A). Then

$$
\begin{aligned}
\mathbf{A}(s, u(s)) = \bigoplus_{m=1}^M A^m(s, u(s)) = \quad & A^1(s, u(s)) \oplus A^1(s, u(s)) \oplus \ldots \oplus A^M(s, u(s)) \\
= \quad & A^1(s, u(s)) \otimes I^{A^2} \otimes \ldots \otimes I^{A^M} \\
+ \quad & I^{A^1} \otimes A^2(s, u(s)) \otimes \ldots \otimes I^{A^M} \\
& \ldots \\
+ \quad & I^{A^1} \otimes I^{A^2} \otimes \ldots \otimes A^M(s, u(s))
\end{aligned}
\tag{5}
$$

is called the tensor sum [12] of the square matrices $A^m(s, u(s))$, $m = 1, \ldots, M$.

Theorem 1. *Let* $\mathbf{X}_t$ *be the global state of the coupled MC system, where each MC satisfies the representation* (1). *Then* $\mathbf{X}_t$ *satisfies the representation*

$$\mathbf{X}_t = \mathbf{X}_0 + \int_0^t \mathbf{A}(s, u(s))\mathbf{X}_s ds + \mathbf{W}_t = \mathbf{X}_0 + \int_0^t \bigoplus_{m=1}^M A^m(s, u(s)) \bigotimes_{m=1}^M X_s^m ds + \mathbf{W}_t, \tag{6}$$

where $\mathbf{W}_t$ *is a square integrable martingale.*

Note that from Theorem 1 it follows, that piece-wise constant right-continuous process $\mathbf{X}_t$ and square integrable martingale $\mathbf{W}_t$ give the solution of the martingale problem (see Reference [13], [Chapter III] for details) for SDE (6).

Proof. The theory of stochastic differential equations with values from finite-dimensional Hilbert spaces can be found in Reference [14], and the present proof is based on the corollary from the Ito formula for a finite-dimensional Hilbert space [14], [Section 4.1]. This result, applied for the case of $M = 2$, that is, for tensor product $\mathbf{X}_t = X_t^1 \otimes X_t^2$, brings the following:

$$X_t^1 \otimes X_t^2 = X_0^1 \otimes X_0^2 + \int_0^t X_{s-}^1 \otimes dX_s^2 + dX_s^1 \otimes X_{s-}^2 + \left[W^1, W^2\right]_t,$$

where $\left[W^1, W^2\right]_t$ stands for the tensor mutual variation of the processes X_t^1, X_t^2, which is a zero tensor due to the independence of the martingales W_t^1, W_t^2.

Substituting the martingale representation of the individual MCs (1) for $m = 1, 2$ we obtain

$$X_t^1 \otimes X_t^2 = X_0^1 \otimes X_0^2 + \int_0^t A^1(s, u(s))X_s^1 \otimes X_s^2 ds + X_s^1 \otimes A^2(s, u(s))X_s^2 ds + \mathbf{W}_t^{1,2},$$

where $d\mathbf{W}_t^{1,2} = X_{s-}^1 \otimes dW_t^2 + dW_t^1 \otimes X_{s-}^2$, so $\mathbf{W}_t^{1,2}$ is a square integrable martingale.

Proceeding further for $m > 2$, we finally have

$$
\begin{aligned}
X_t^1 \otimes X_t^2 \otimes \ldots \otimes X_t^M = \quad & X_0^1 \otimes X_0^2 \otimes \ldots \otimes X_0^M \\
+ & \int_0^t A^1(s, u(s))X_s^1 \otimes X_s^2 \otimes \ldots \otimes X_s^M ds \\
+ & \quad X_s^1 \otimes A^2(s, u(s))X_s^2 \otimes \ldots \otimes X_s^M ds \\
& \cdots \\
+ & \quad X_s^1 \otimes X_s^2 \otimes \ldots \otimes A^M(s, u(s))X_s^M ds \quad + \mathbf{W}_t.
\end{aligned}
\tag{7}
$$

Now using the definition of tensor product of matrices (see Appendix A), we can factorize the summands in the last expression as

$$
\begin{aligned}
X_t^1 \otimes X_t^2 \otimes \ldots \otimes X_t^M = \quad & X_0^1 \otimes X_0^2 \otimes \ldots \otimes X_0^M \\
+ & \int_0^t \left(A^1(s, u(s)) \otimes I^{A^2} \otimes \ldots \otimes I^{A^M}\right) \left(X_s^1 \otimes X_s^2 \otimes \ldots \otimes X_s^M\right) ds \\
+ & \quad \left(I^{A^1} \otimes A^2(s, u(s)) \otimes \ldots \otimes I^{A^M}\right) \left(X_s^1 \otimes X_s^2 \otimes \ldots \otimes X_s^M\right) ds \\
& \cdots \\
+ & \quad \left(I^{A^1} \otimes I^{A^2} \otimes \ldots \otimes A^M(s, u(s))\right) \left(X_s^1 \otimes X_s^2 \otimes \ldots \otimes X_s^M\right) ds \quad + \mathbf{W}_t,
\end{aligned}
$$

which, along with the definition of the tensor sum (5), yields the theorem statement. □

Note that Formula (7) provides an efficient way to calculate the tensor contraction $\mathbf{A}(s, u(s))\mathbf{X}_s = \mathbf{A}^{i^1...i^M}_{j^1...j^M}\mathbf{X}_s^{j^1...j^M}$:

$$
\begin{aligned}
\mathbf{A}(s, u(s))\mathbf{X}_s = \quad & A^1(s, u(s))X_s^1 \otimes X_s^2 \otimes \ldots \otimes X_s^M \\
+ \quad & X_s^1 \otimes A^2(s, u(s))X_s^2 \otimes \ldots \otimes X_s^M \\
& \ldots \\
+ \quad & X_s^1 \otimes X_s^2 \otimes \ldots \otimes A^M(s, u(s))X_s^M.
\end{aligned}
\tag{8}
$$

4. Optimal Control

4.1. Value Function Representation

The value function of the coupled MC system is a function, which is equal to the minimum total cost for the system given the starting time $t \in [0, T]$ and state $\mathbf{X}_t = \mathbf{X}$:

$$
V(t, \mathbf{X}) = \inf_{u(\cdot)} J[u(\cdot), \mathbf{X}(\cdot) \mid \mathbf{X}_t = \mathbf{X}],
$$

where

$$
J[u(\cdot), \mathbf{X}(\cdot) \mid \mathbf{X}_t = \mathbf{X}] = \mathbb{E}\left[\langle \boldsymbol{\varphi}_T, \mathbf{X}_T \rangle + \int_t^T \langle \mathbf{f}(s, u(s)), \mathbf{X}_s \rangle ds \,\middle|\, \mathbf{X}_t = \mathbf{X} \right].
$$

As in Formula (4) for the running cost function, we now represent the value function $V(t, \mathbf{X})$ as

$$
V(t, \mathbf{X}) = \langle \boldsymbol{\varphi}_t, \mathbf{X} \rangle,
\tag{9}
$$

where $\boldsymbol{\varphi}_t$ is a multidimensional array-valued function defined for any possible global state of the coupled MC system $\mathbf{X}_t = \mathbf{e}^{i^1...i^M} = e^{i^1} \otimes \ldots \otimes e^{i^M}$.

4.2. Dynamic Programming Equation

Further, we generalize the approach to the control optimization of ordinary MC [7] to the controlled MC systems. Define the *dynamic programming equation* with respect to $\hat{\boldsymbol{\varphi}}_t \in \mathbb{R}^{N^1} \times \ldots \times \mathbb{R}^{N^M}$:

$$
\frac{\langle d\hat{\boldsymbol{\varphi}}_t, \mathbf{X}_t \rangle}{dt} = -\min_{u \in U}\left\{ \langle \hat{\boldsymbol{\varphi}}_t, \mathbf{A}(t, u)\mathbf{X}_t \rangle + \langle \mathbf{f}(t, u), \mathbf{X}_t \rangle \right\} = -\min_{u \in U} H(t, \hat{\boldsymbol{\varphi}}_t, u, \mathbf{X}_t) = -\mathcal{H}(t, \hat{\boldsymbol{\varphi}}_t, \mathbf{X}_t)
\tag{10}
$$

with terminal condition $\hat{\boldsymbol{\varphi}}(T) = \boldsymbol{\varphi}_T$ [7,10,15]. Our aim is to show that (10) has a unique solution $\hat{\boldsymbol{\varphi}}_t$, and that $\hat{\boldsymbol{\varphi}}_t = \boldsymbol{\varphi}_t$ from the value function representation (9). This will allow us to present an optimal Markov control, which minimizes the performance criterion (3).

If Assumption 1 holds, then $H(t, \hat{\boldsymbol{\varphi}}_t, u, \mathbf{X})$ is continuous in (t, u), and hence for any $\mathbf{X} \in \mathbf{S}$ there exists

$$
u^*(t, \mathbf{X}) = \underset{u \in U}{\operatorname{argmin}}\, H(t, \hat{\boldsymbol{\varphi}}_t, u, \mathbf{X}),
$$

which minimizes this continuous function on a compact set U. Moreover, $H(t, \hat{\boldsymbol{\varphi}}_t, u, \mathbf{X})$ is affine in $\hat{\boldsymbol{\varphi}}_t$ for any $(t, \mathbf{X}) \in [0, T] \times \mathbf{S}$, so $\mathcal{H}(t, \hat{\boldsymbol{\varphi}}_t, \mathbf{X})$ is Lipschitz in $\hat{\boldsymbol{\varphi}}_t$, and hence Equation (10) has a unique solution on $[0, T]$.

The existence of an $\mathcal{F}_t^{\mathbf{X}}$-predictable (i.e., $\mathcal{F}_{t-}^{\mathbf{X}}$-measurable) optimal control, provided that Assumption 1 is true, follows from a general result [16], [Theorem 4.2]. The following theorem states that the optimal control can be chosen among the Markov controls, which depend not on the whole process history $\mathbf{X}|_0^t$, but on the left limit $\mathbf{X}_{t-}$ only. It should be noted that this result is a generalization of Theorem 2.8 [7] for controlled MCs to the case of controlled MC systems.

Theorem 2. *If Assumption 1 holds, then the Markov control and the value function is defined by the solution of the DPE (10), that is, $V(t, \mathbf{X}) = \langle \hat{\boldsymbol{\varphi}}_t, \mathbf{X} \rangle$ and*

$$\hat{u}(t, \mathbf{X}|_0^t) = u^*(t, \mathbf{X}_{t-}) = \underset{u \in U}{\operatorname{argmin}} \, H(t, \hat{\boldsymbol{\varphi}}_t, u, \mathbf{X}_{t-}),$$

is the optimal control.

Proof. Since $\hat{u}(\cdot)$ is a predictable control, then for any initial condition $\mathbf{X}_0 \in \mathbf{S}$ there exists a unique solution of the martingale problem (6). That is, there exists a process $\mathbf{X}_t^{\hat{u}} \in SD_{[0,T]}$, where $SD_{[0,T]}$ is a class of **S**-valued right continuous functions with left limits, which satisfies Equation (6), that is,

$$\mathbf{X}_t^{\hat{u}} = \mathbf{X}_0 + \int\limits_0^t \mathbf{A}(s, \hat{u}(s)) \mathbf{X}_s^{\hat{u}} ds + \mathbf{W}_t^{\hat{u}},$$

where $\mathbf{W}^{\hat{u}}$ is a square integrable $\mathcal{F}_t^{\mathbf{X}^{\hat{u}}}$ martingale.

Let us take some admissible control $u(\cdot)$ and corresponding solution $\mathbf{X}_t^u$ of the martingale problem (6). Applying Ito's formula to the process $\langle \hat{\boldsymbol{\varphi}}_t, \mathbf{X}_t^u \rangle$ we get

$$\langle \hat{\boldsymbol{\varphi}}_t, \mathbf{X}_t^u \rangle = \langle \hat{\boldsymbol{\varphi}}_T, \mathbf{X}_T^u \rangle - \int\limits_t^T \left\{ \langle d\hat{\boldsymbol{\varphi}}_s, \mathbf{X}_s^u \rangle + \langle \hat{\boldsymbol{\varphi}}_s, d\mathbf{X}_s^u \rangle \right\}.$$

Since the first integrand in the RHS is the minimum in (10), and $\mathbf{X}_s^u$ in the second integrand obeys the martingale representation (6), we can transform the last equation into the following inequality:

$$\begin{aligned}
\langle \hat{\boldsymbol{\varphi}}_t, \mathbf{X}_t^u \rangle &\leq \langle \hat{\boldsymbol{\varphi}}_T, \mathbf{X}_T^u \rangle + \int\limits_t^T \left\{ H(t, \hat{\boldsymbol{\varphi}}_s, u(s), \mathbf{X}_s^u) ds - \langle \hat{\boldsymbol{\varphi}}_s, \mathbf{A}(s, u(s)) \mathbf{X}_t^u \rangle ds - \langle \hat{\boldsymbol{\varphi}}_s, d\mathbf{W}_t^u \rangle \right\} \\
&= \langle \hat{\boldsymbol{\varphi}}_T, \mathbf{X}_T^u \rangle + \int\limits_t^T \left\{ \langle \mathbf{f}(s, u(s)) \mathbf{X}_t^u \rangle ds - \langle \hat{\boldsymbol{\varphi}}_s, d\mathbf{W}_t^u \rangle \right\}.
\end{aligned}$$

Taking the expectation $\mathbb{E}\left[\cdot \mid \mathbf{X}_t = \mathbf{X} \right]$ of both parts of the inequality, we get

$$J[u(\cdot), \mathbf{X}(\cdot) \mid \mathbf{X}_t = \mathbf{X}] = \mathbb{E}\left[\langle \hat{\boldsymbol{\varphi}}_T, \mathbf{X}_T^u \rangle + \int\limits_t^T \langle \mathbf{f}(s, u(s)), \mathbf{X}_s^u \rangle ds \,\middle|\, \mathbf{X}_t = \mathbf{X} \right] \geq \langle \hat{\boldsymbol{\varphi}}_t, \mathbf{X} \rangle,$$

since $\hat{\boldsymbol{\varphi}}_s$ is a continuous deterministic function and the expectation of the integral over the martingale is equal to zero. Substitution of the control $\hat{u}(t, \mathbf{X}|_0^t) = u^*(t, \mathbf{X}_{t-})$ yields the equality, which means that the solution to the DPE (10) defines the value function $V(t, \mathbf{X}) = \langle \hat{\boldsymbol{\varphi}}_t, \mathbf{X} \rangle$ and $u^*(t, \mathbf{X}_{t-})$ is the optimal control. $\square$

4.3. Optimal Control Calculation

In the previous section it was shown that the solution to the DPE (10) is unique and from Theorem 2 it follows that the Markov control, which minimizes the RHS of this equation, is optimal. Now if we let $\mathbf{X}_t = \mathbf{e}^{i^1 \dots i^M} = e^{i^1} \otimes \dots \otimes e^{i^M}$, then we get a system of $\prod_{m=1}^{M} N^m$ ordinary differential equations (ODEs):

$$\frac{d\hat{\boldsymbol{\varphi}}_t^{i^1 \dots i^M}}{dt} = -\min_{u \in U} \left\{ \langle \hat{\boldsymbol{\varphi}}_t, \mathbf{A}(t, u) \mathbf{e}^{i^1 \dots i^M} \rangle + \mathbf{f}_{i^1 \dots i^M}(t, u) \right\} = -\mathcal{H}(t, \hat{\boldsymbol{\varphi}}_t, \mathbf{e}^{i^1 \dots i^M}). \tag{11}$$

The inner product in the right-hand side of the ODEs (11) can be simplified as follows:

$$\langle \hat{\boldsymbol{\varphi}}_t, \mathbf{A}(t,u)\mathbf{e}^{i^1\dots i^M}\rangle = \sum_{j^1\dots j^M} \hat{\varphi}_t^{j^1\dots j^M} \langle \mathbf{A}(t,u)\mathbf{e}^{i^1\dots i^M}, \mathbf{e}^{j^1\dots j^M}\rangle$$

$$= \sum_{j^1\dots j^M} \hat{\varphi}_t^{j^1\dots j^M} \left(\langle A^1(t,u(t))e^{i^1}, e^{j^1}\rangle \otimes \dots \otimes \langle e^{i^M}, e^{j^M}\rangle + \dots + \langle e^{i^1}, e^{j^1}\rangle \otimes \dots \otimes \langle A^M(t,u(t))e^{i^M}, e^{j^M}\rangle \right)$$

$$= \sum_{j^1\dots j^M} \hat{\varphi}_t^{j^1\dots j^M} \left(\langle A^{1^T}(t,u(t))e^{j^1}, e^{i^1}\rangle \otimes \dots \otimes \langle e^{j^M}, e^{i^M}\rangle + \dots + \langle e^{j^1}, e^{i^1}\rangle \otimes \dots \otimes \langle A^{M^T}(t,u(t))e^{j^M}, e^{i^M}\rangle \right)$$

$$= \sum_{j^1\dots j^M} \hat{\varphi}_t^{j^1\dots j^M} \langle \mathbf{A}^T(t,u(t))\mathbf{e}^{j^1\dots j^M}, \mathbf{e}^{i^1\dots i^M}\rangle = \langle \mathbf{A}^T(t,u(t))\hat{\boldsymbol{\varphi}}_t, \mathbf{e}^{i^1\dots i^M}\rangle,$$

where we used the definition of tensor sum (5), the representation of the inner product as a component-wise sum, and denoted

$$\mathbf{A}^T(t,u(t)) = \bigoplus_{m=1}^{M} A^{m^T}(t,u(t)).$$

Finally, we get the following tensor form for the system of ODEs (11):

$$\frac{d\hat{\boldsymbol{\varphi}}_t}{dt} = -\min_{u\in U}\left\{ \mathbf{A}^T(t,u(t))\hat{\boldsymbol{\varphi}}_t + \mathbf{f}(t,u) \right\}. \tag{12}$$

The system (12) provides a method of simultaneous calculation of the optimal Markov control $u^*(t,\mathbf{X}_{t-})$ and the DPE solution. This system is solved backwards in time, starting from the terminal condition $\hat{\boldsymbol{\varphi}}(T) = \hat{\boldsymbol{\varphi}}(t)$. The minimizations of the RHSs of the equations of this system are independent of each other and, in the case of the numerical solution, these time-consuming operations can be performed in parallel. The result of the minimization yields a set of control functions $u^*_{i^1\dots i^M}(t)$, which define the Markov optimal control:

$$u^*(t,\mathbf{X}_{t-}) = u^*_{i^1\dots i^M}(t)I\{\mathbf{X}_{t-} = \mathbf{e}^{i^1\dots i^M}\}. \tag{13}$$

4.4. Kolmogorov Equation

Once the optimal control is obtained, there arises the problem of the state probabilities calculation. The main difficulty is that the optimal Markov control depends on the state $\mathbf{X}_t$, so one cannot directly compute the generator using the tensor sum Formula (5). To overcome this issue for a single controlled MC with the generator $A(t,u(t,X_t))$, it suffices to substitute $u(t,X_t)$ in the j-th column with control $u(t,e^j)$, which corresponds to the state e^j. For the controlled MC system, the generator $\mathbf{A}(t,u^*(t,\mathbf{X}_t))$ can be constructed in a similar block-wise manner. Indeed, for any global state $\mathbf{e}^{j^1\dots j^M} = e^{j^1} \otimes \dots \otimes e^{j^M}$, the corresponding Markov control is equal to $u^*(t,\mathbf{e}^{j^1\dots j^M})$ and

$$\mathbf{A}^{i^1\dots i^M}_{j^1\dots j^M}(t,u^*(t,\mathbf{X}_t)) = \mathbf{A}^{i^1\dots i^M}_{j^1\dots j^M}(t,u^*(t,\mathbf{e}^{j^1\dots j^M})), \quad \text{for every multi-index } i^1\dots i^M, \tag{14}$$

where $\mathbf{A}(t,u^*(t,\mathbf{e}^{j^1\dots j^M}))$ is defined by Formula (5) for $u(t) = u^*(t,\mathbf{e}^{j^1\dots j^M})$.

Finally, Equation (14) defines the transition rate tensor $\mathbf{A}(t,u^*(t,\mathbf{X}_t))$, which is nonrandom, and the state probabilities $\mathbf{P}_t = \mathbb{P}\{\mathbf{X}_t = \mathbf{e}^{i^1\dots i^M}\}$ for the process $\mathbf{X}_t^{\hat{u}}$ satisfy the Kolmogorov equation

$$\frac{d\mathbf{P}_t}{dt} = \mathbf{A}(t,u^*(t,\mathbf{X}_t))\mathbf{P}_t. \tag{15}$$

Note that as in the case of a single controlled MC this equation is simply a result of taking the expectation of both sides of the corresponding SDE (6). Note also that the right-hand side of (15) is the tensor contraction $\mathbf{A}(t,u^*(t,\mathbf{X}_t))\mathbf{P}_t = \mathbf{A}^{i^1\dots i^M}_{j^1\dots j^M}(t,u^*(t,\mathbf{X}_t))\,\mathbf{P}_t^{j^1\dots j^M}$ defined as in (8).

5. Numerical Study: Goulburn River Cascade

To illustrate our approach to the modeling of real-world systems as coupled MCs and optimal control synthesis, in this section we present a numerical study based on real-world data. The usual area of applications for MCs and numerous generalizations is service systems, which can be described in terms of Poisson arrival flows, queues, and servers. For example, in References [6,17], one can find an application of coupled MCs for congestion control in networks. Specifically, an optimal control strategy was proposed for a multi-homing network connection, where the packets (portions of incoming traffic) could be sent through one of two lines of different speed and price or discarded. However, in the present paper, we stick to another application area, namely, the control of water reservoirs. The reason is that the proposed approach allows successful modeling of these systems, accounting for the stochastic nature of the incoming and outgoing flows, being dependent on the weather condition and users' behavior. Moreover, the proposed form of the performance criterion (3) permits a variety of optimization goals, for example, minimization of the difference between the actual and desired demand intensities, maintenance of the balance between the incoming and outgoing flows, minimization of the probability that either dam falls below some critical level on average or at the terminal time [9]. Besides, this criterion allows penalizing control in certain states, so that natural restrictions (such as the impossibility to transfer water from an empty or to an overflowing reservoir) can also be taken into account. The terminal condition can reflect the desire to bring the system to a certain state by the end of the control interval.

The object of our study is a system of interconnected dams of the Goulburn River cascade. The model is deliberately simplified to avoid cumbersome details, which are unnecessary for a journal on mathematical subjects, nevertheless, it demonstrates all the necessary techniques, such as MCs states' choice, relations between the discharges and state transition intensities, accounting for the natural inflows and outflows, consumer demands, and environmental constraints.

The Goulburn River Basin is located in the south-western part of State Victoria, Australia. The description of the region and a review of the literature devoted to water resource modeling and planning can be found, for instance, in References [18–20]. The review of papers on allocation modeling in this region and water trading can be found in References [21–23]. The Goulburn Basin, together with the Upper Murray, Ovens, Kiewa, Broken, Loddon, and Campaspe Basins constitute an area called the Goulburn-Murray Irrigation District (common abbreviation is GMW from Goulburn-Murray Water—an organization governing water resources in this area). Since the Goulburn River is a major waterway and a major contributor of water supply in the district, it is an object of very intensive research and modeling. The GMW uses the Goulburn simulation model (GSM) for the optimization and planning of water supply in the region. The GSM represents the Goulburn system as a set of nodes of different types: storages, demands, and streamflow inputs. These nodes are connected by a network of carriers characterized by their capacities and delivery penalties (penalty functions). The GSM was calibrated using a trial-and-error procedure, which optimizes the model parameters related to the water supply infrastructure [24]. The model presented here reflects a part of the GSM and involves Lake Eildon—a major storage of water, which can be required by farmers during the irrigation season, and Lake Nagambie with the Goulbourn Weir, which is connected by three major channels to the end-users. The Goulburn River connects the two water storages and continues to run after Nagambie, serving as its fourth outflow. Based on the request of irrigators, the water is released from Lake Eildon to Lake Nagambie, and then by one of three channels or the river delivered to the farmers (water right holders). The important point is that apart from irrigation requests, the system must satisfy environmental demand, which means the levels in rivers and channels should exceed some minimal threshold that will allow the support of healthy ecosystems and provide some required minimal inflows to satisfy downstream demands.

The scheme of the Goulburn River cascade is presented in Figure 1. We assume that the water level of Lake Eildon is affected by two major flows. One is the incoming flow, which reflects the sum of all the natural water arrivals and losses, including precipitation, upstream tributaries,

evaporation, and other causes. The single outcoming flow is the controlled discharge of the Goulburn River, which also serves as the single incoming flow of Lake Nagambie (other natural inflows and outflows into this lake are ignored here due to a significant difference in the size of the two lakes). From Lake Nagambie, there are four controlled discharges, namely the Stuart Murray Canal, Cattanach Canal, East Goulburn Main Channel, and the Goulburn River.
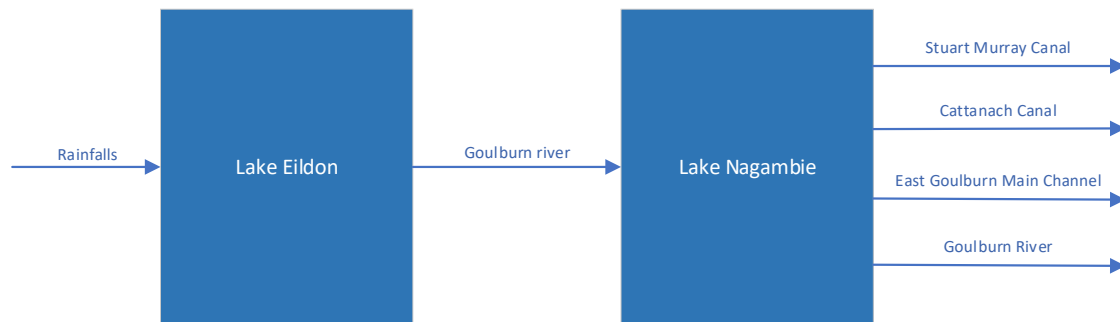

Figure 1. Goulburn River cascade scheme.

5.1. Mcs and Generators

The states of the water reservoirs indicate their water levels, expressed in equal portions of their volume. So for Lake Eildon with a total volume equal to $V_E = 3{,}390{,}000$ ML, the bounds of the states are $\frac{i \cdot V_E}{N_E}$, where $i = 1, \ldots, N_E$ and N_E is the desired number of states, which is chosen according to accuracy demands. The same state division is considered for Lake Nagambie with volume $V_N = 25{,}500$ ML: the bounds are $\frac{j \cdot V_N}{N_N}$, $j = 1, \ldots, N_N$, N_N — desired number of states for this reservoir. Association of the states with water volumes instead of the water surface levels leads to the same dependency of the transitions on the incoming and outcoming flows and hence simplifies the resulting model.

We model the reservoirs as MCs, so the evolution of the reservoirs' states in time is described by the stochastic differential Equation (1), where $M = 2$ and X_t^1 and X_t^2 stand for Lakes Eildon and Nagambie respectively. Due to natural reasons, the transitions are only possible between adjacent states, so the generators $A^m(t, u)$ reflect a birth-and-death processes:

$$A^1(t, u) = \begin{pmatrix} -L_R & \lambda_R & 0 & 0 & \ldots & 0 \\ U_E & -U_E - L_R & L_R & 0 & \ldots & 0 \\ 0 & U_E & -U_E - L_R & L_R & \ldots & 0 \\ & & \ldots & & & \\ 0 & \ldots & 0 & 0 & U_E & -U_E, \end{pmatrix}$$

where $L_R = L_R(t)$ is the intensity of the incoming flow of natural water arrivals and losses, and $U_E = U_E(t)$ is the controlled intensity, which reflects the flow from Lake Eildon to Lake Nagambie. The variables L_R and U_E have a purely probabilistic sense, which has to be translated into the common units of water discharge measurement, say $\left[\frac{ML}{day}\right]$. Here the basic assumption is the equality of the average transition times, which for the constant rate λ and constant discharge u are equal to $1/\lambda$ from the one side, and to $\frac{v}{u}$ from the other. Here v is the volume of water, which corresponds to the difference between the levels, for example, for Lake Eildon, $v = \frac{V_E}{N_E}$. Thus, in terms of the discharges, the generator $A^1(t, u)$ can be represented as follows:

$$A^1(t, u) = \frac{N_E}{V_E} \begin{pmatrix} -\lambda_R & \lambda_R & 0 & 0 & \ldots & 0 \\ u_E & -u_E - \lambda_R & \lambda_R & 0 & \ldots & 0 \\ 0 & u_E & -u_E - \lambda_R & \lambda_R & \ldots & 0 \\ & & \ldots & & & \\ 0 & \ldots & 0 & 0 & u_E & -u_E, \end{pmatrix} \tag{16}$$

where u_E is the discharge from Lake Eildon, and λ_R is the sum of all the natural inflows expressed in common units $\left[\frac{ML}{day}\right]$.

Denote the controlled discharges from Lake Nagambie by u_{SM}, u_{Ca}, u_{EG}, u_{GR}, which stand respectively for the discharges of the Stuart Murray, Cattanach, East Goulburn Main Channels, and the Goulburn River, and the sum of these discharges as $u_N = u_{SM} + u_{Ca} + u_{EG} + u_{GR}$. Since the outflow of Lake Eildon is at the same time the incoming flow of Lake Nagambie, we have the following representation of the generator $A^2(t, u)$:

$$A^2(t, u) = \frac{N_N}{V_N} \begin{pmatrix} -u_E & u_E & 0 & 0 & \dots & 0 \\ u_N & -u_N - u_E & u_E & 0 & \dots & 0 \\ 0 & u_N & -u_N - u_E & u_E & \dots & 0 \\ & & \dots & & & \\ 0 & & \dots & 0 & 0 & u_N & -u_N. \end{pmatrix} \tag{17}$$

It should be noted that the capacity of Lake Eildon is more than 100 times larger than the capacity of Lake Nagambie. This causes certain problems for the numerical simulation since the same discharge values affect the transition rates in different scales. To overcome this issue, one can normalize the transition rates by adjusting the number of states, providing $N_E \gg N_N$. Another way is to scale down the time discretization step, ensuring adequate probabilities of state transitions for the MC, which corresponds to Lake Nagambie.

5.2. Performance Criterion and Constraints

The performance criterion should take into account the customer demands and ensure the balance of the water reservoirs inflows and outflows. We assume that there are some reference values for the demands in the channels and the Goulburn River: $\tilde{u}_{SM}$, $\tilde{u}_{Ca}$, $\tilde{u}_{EG}$, $\tilde{u}_{GR}$, which an optimal control aims to satisfy. Let

$$\begin{aligned} \mathbf{f}_{ij}(t, u) = \quad & (u_N - u_E)^2 \\ + \quad & (u_{SM} - \tilde{u}_{SM})^2 + (u_{Ca} - \tilde{u}_{Ca})^2 + (u_{EG} - \tilde{u}_{EG})^2 + (u_{GR} - \tilde{u}_{GR})^2 \\ + \quad & Cu_E^2 I\{i = 1\} + Cu_{SM}^2 I\{j = 1\} + Cu_{Ca}^2 I\{j = 1\} + Cu_{EG}^2 I\{j = 1\} + Cu_{GR}^2 I\{j = 1\}, \end{aligned} \tag{18}$$

where i and j stand for the state number of Lakes Eildon and Nagambie respectively, C is some large number, and $I\{\cdot\}$ is an indicator function.

The running cost function (4) with the state-related cost given by (18) reflects the mentioned aim along with the intention to balance the arriving and outgoing flows of Lake Nagambie. Plus, this running cost penalizes the nonzero discharges from both lakes, when they are in the states $X_t^1 = e_1$, $X_t^2 = e_1$, which corresponds to the lowest possible level, or drought.

The terminal condition $\varphi_0(\mathbf{X}_T) = \langle \varphi_T, \mathbf{X}_T \rangle$ in (3) is defined by matrix φ_T with elements φ_T^{ij}, which correspond to the combination of states $X_T^1 = i$ and $X_T^2 = j$. Let $\tilde{\mathbf{X}}_T$ be the desired terminal state of the system, then define the terminal condition as follows:

$$\varphi_T^{ij} = C \cdot \exp(\|(i, j)^T - \tilde{\mathbf{X}}_T\|_1), \tag{19}$$

where $\| \cdot \|_1$ is the L^1 norm (sum of absolute values). This condition penalizes any deviation of the system's state at the end of the control interval T from the desired and, moreover, it assigns a larger penalty for greater deviations.

For all the control variables we also define phase constraints in the form of upper and lower bounds:

$$\underline{u} \le u \le \bar{u}, \quad u = (u_E, u_{SM}, u_{Ca}, u_{EG}, u_{GR})^T. \tag{20}$$

These constraints bound the maximal possible discharges for the channels and the Goulburn River and guarantee the environmental demand, mentioned earlier.

5.3. Parameters Definition

In previous subsections, the Goulburn River cascade model was defined in general. Nevertheless, there still are some undefined parameters, which will be specified using real-world data. To define the reference demand values $\tilde{u}_{SM}$, $\tilde{u}_{Ca}$, $\tilde{u}_{EG}$, $\tilde{u}_{GR}$, we use the data on the discharges of the corresponding channels and river available from the Water Measurement Information System of the Department of Environment, Land, Water & Planning of Victoria State Government [25]. The data on the daily discharges was grouped by the day of year, and for each group, the mean value was calculated. The resulting time series were approximated by functions

$$u(t) = a_0 + \sum_{i=1}^{K} \left\{ a_i \sin\left(\frac{2\pi i}{T}t\right) + b_i \cos\left(\frac{2\pi i}{T}t\right) \right\},$$

where $T = 365$ days and coefficients a_i, b_i are chosen using the Least-Squares method. The resulting functions are smooth and have a period equal to one year, which makes them suitable for the reference values of the irrigation demands in the proposed model. The time series of the average daily discharges and the approximating functions are presented in Figure 2.
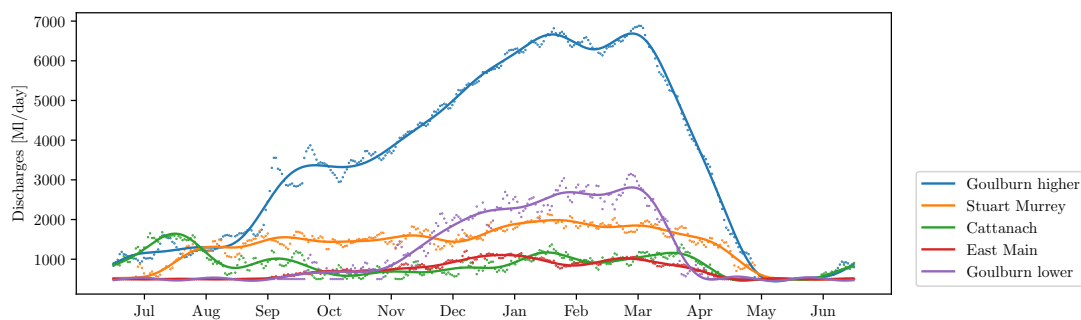
Figure 2. Average daily discharges (**dotted lines**) and the approximating smooth periodic functions (**solid lines**).

The data for the Lake Eildon incoming flow is obtained from the Climate Data Online service of the Bureau of Meteorology of the Australian Government [26]. The rainfall data is collected at the weather station situated there. This data is averaged on a daily basis and approximated by a smooth periodic function in the same way as the daily discharges. The resulting function serves as an incoming flow pattern $\Lambda_R(t)$. For this model, we assume that the irrigation demands agree with the total available natural resources, so the incoming flow pattern is normalized to make the incoming flow $\lambda_R(t) = const \cdot \Lambda_R(t)$ satisfy the following condition:

$$\int_0^T (\tilde{u}_{SM}(t) + \tilde{u}_{Ca}(t) + \tilde{u}_{EG}(t) + \tilde{u}_{GR}(t))dt = \int_0^T \lambda_R(t)dt.$$

The resulting flow approximation along with normalized rainfall averages is presented in Figure 3.
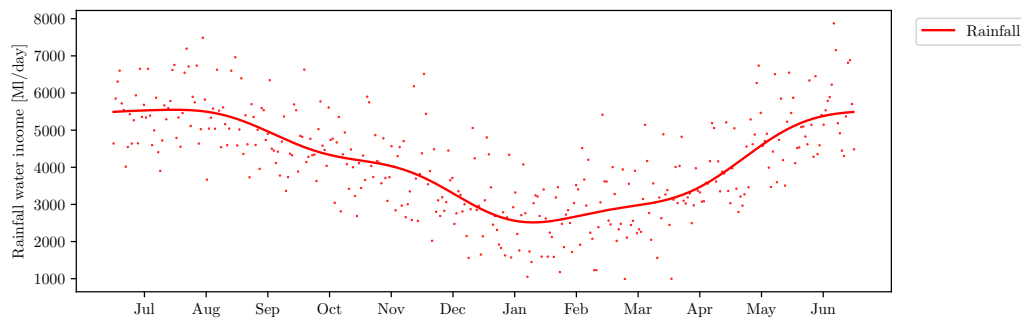
Figure 3. Normalized rainfall daily averages (dotted line) and the approximation (solid line).

Specifying of the upper and lower bounds for the control parameters $\underline{u} = (\underline{u}_E, \underline{u}_{SM}, \underline{u}_{Ca}, \underline{u}_{EG}, \underline{u}_{GR})^T$ and $\bar{u} = (\bar{u}_E, \bar{u}_{SM}, \bar{u}_{Ca}, \bar{u}_{EG}, \bar{u}_{GR})^T$ by maximum and minimum registered values of the corresponding discharges time series obtained from Reference [25] finalizes the definition of our Goulburn River cascade model.

5.4. Simulation Results

The complete definition of the control optimization problem is given by

- state Equation (1) with generators (16) and (17);
- optimization criterion (3) with running cost (18) and terminal condition (19);
- phase constraints (20);
- incoming flow $\lambda_R(t)$, reference values of the irrigation demands $\tilde{u}_{SM}$, $\tilde{u}_{Ca}$, $\tilde{u}_{EG}$, $\tilde{u}_{GR}$ and upper and lower bounds for the control parameters defined in Section 5.3.

Now the optimal Markov control (13) can be calculated as a solution to the system of ODEs (12). Remember that the minimization in the RHS of the ODEs can be done independently and hence can be effectively parallelized. The system (12) was solved numerically using the Euler method with the discretization step $\delta t = 10^{-4}$, while the time scale and all the necessary functions were normalized so that $T = 1.0$. The number of MCs' states for both lakes was chosen as $N_E = N_N = 10$, and the desired terminal state $\tilde{\mathbf{X}}_T = (9,9)^T$.

In Figures 4 and 5, the optimal control values are presented for the states $(9,9)^T$ and $(7,9)^T$. One can see the nontrivial behavior of the optimal discharges, which differs from the reference values even for the desired terminal state. The evolution of the value function components φ_t^{ij} exhibits the influence both from the integral part of the criterion and the terminal condition. Indeed, for $\tilde{\mathbf{X}}_T = (9,9)^T$ it monotonically decreases since the demands here are mostly satisfied and the system is already in the desired state. For the state $(7,9)^T$, on the contrary, there are intervals of monotonic growth: the first is because of insufficient supplies in late summer and autumn months, and the second starts in spring and shows the necessity to transit to the desired terminal state $\tilde{\mathbf{X}}_T = (9,9)^T$.

The Kolmogorov Equation (15) allows us to calculate the probabilities of the global system states given the initial state or distribution. The probabilities for states $(9,9)^T$ and $(7,9)^T$ are presented in Figures 4 and 5 on the upper subplots. The initial state here was chosen to be equal to the terminal $\tilde{\mathbf{X}}_0 = \tilde{\mathbf{X}}_T = (9,9)^T$. The solid red lines present the solution of (15), and the dotted lines are the result of Monte-Carlo sampling: the state rate was estimated using a set of 100 sample paths governed by the optimal control strategy. The state probabilities allow us to also calculate the average states of MCs, which can be converted into the average levels of Lakes Eildon and Nagambie. In Figure 6, the average levels calculated by means of the Kolmogorov Equation (15) are presented by solid lines. The corresponding averages calculated using the Monte-Carlo sampling are depicted with dotted lines.
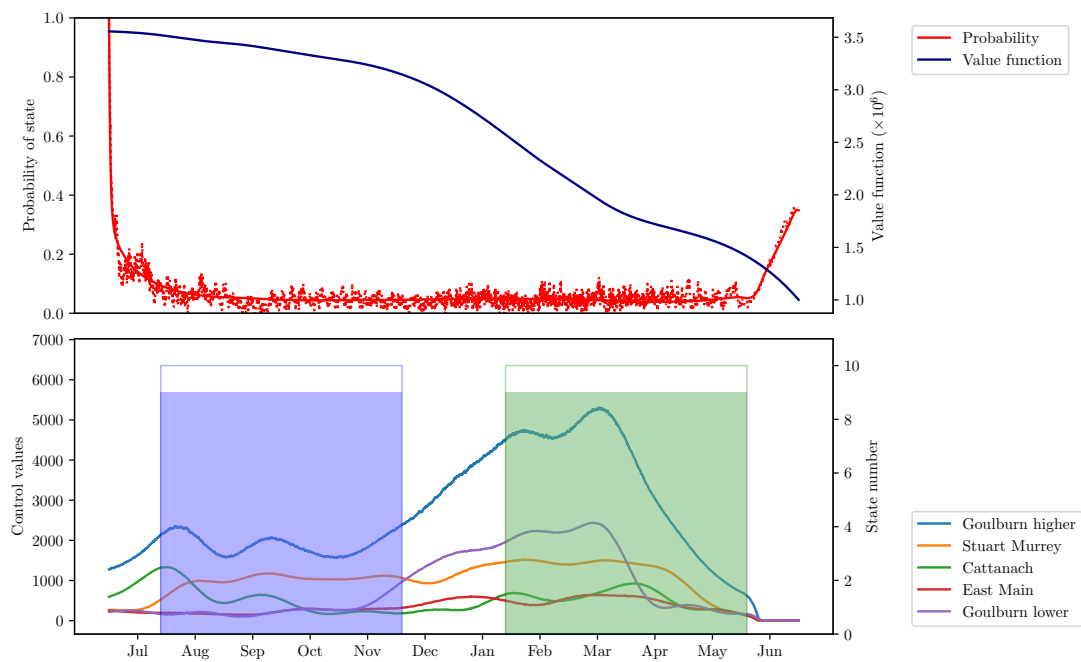
Figure 4. Probability of the state $X^1(t) = 9$, $X^2(t) = 9$, corresponding value function component $\boldsymbol{\varphi}_t^{9,9}$ and optimal Markov control $u_{9,9}^*$.
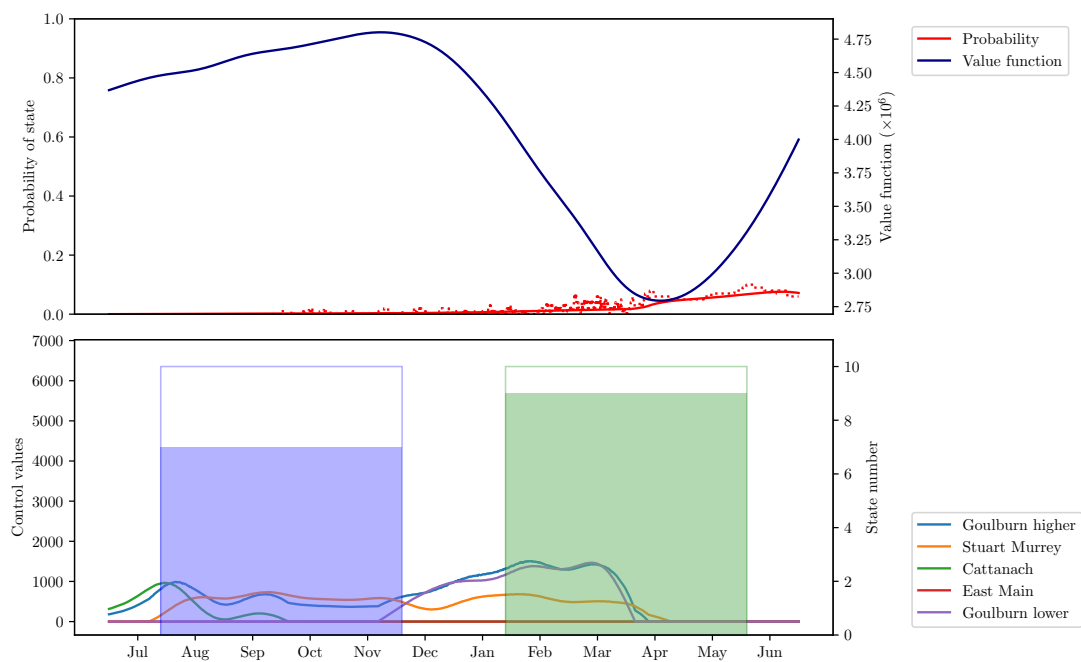
Figure 5. Probability of the state $X^1(t) = 7$, $X^2(t) = 9$, corresponding value function component $\boldsymbol{\varphi}_t^{7,9}$ and optimal Markov control $u_{7,9}^*$.
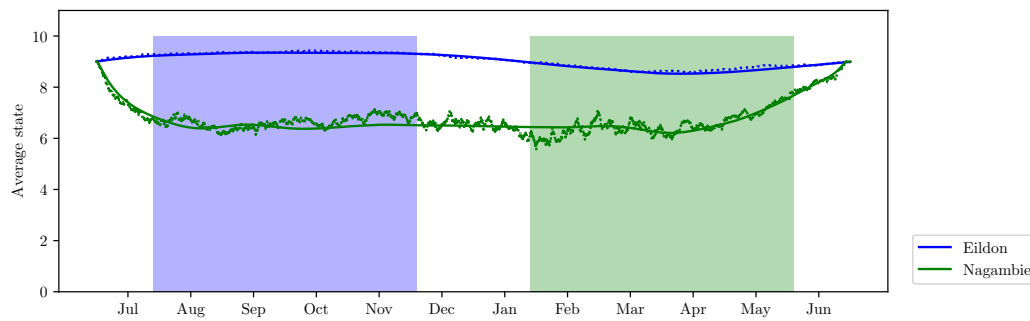
Figure 6. Average levels of Lakes Eildon and Nagambie for optimal control.

As an alternative to the optimal control $u^*(t, \mathbf{X}_t)$ we present here a program control $u^p(t, \mathbf{X}_t)$ which aims to fully satisfy the customer demands:

$$
u^p(t, \mathbf{X}_t) = \begin{pmatrix} I\{\mathbf{X}_t = e^1 \otimes e^j\}(\tilde{u}_{SM} + \tilde{u}_{Ca} + \tilde{u}_{EG} + \tilde{u}_{GR}) \\ I\{\mathbf{X}_t = e^i \otimes e^1\}\tilde{u}_{SM} \\ I\{\mathbf{X}_t = e^i \otimes e^1\}\tilde{u}_{Ca} \\ I\{\mathbf{X}_t = e^i \otimes e^1\}\tilde{u}_{EG} \\ I\{\mathbf{X}_t = e^i \otimes e^1\}\tilde{u}_{GR} \end{pmatrix}.
$$

The indicator functions make sure that the natural constraints are satisfied, that is, the water cannot be taken from an empty dam. In Figure 7, the average levels calculated by means of the Kolmogorov equation and Monte-Carlo sampling are presented. One can observe that this program control leads to overuse of the resources in contrast with the optimal one, which is able to refill the dams after the winter demand growth and rainfall shortage.
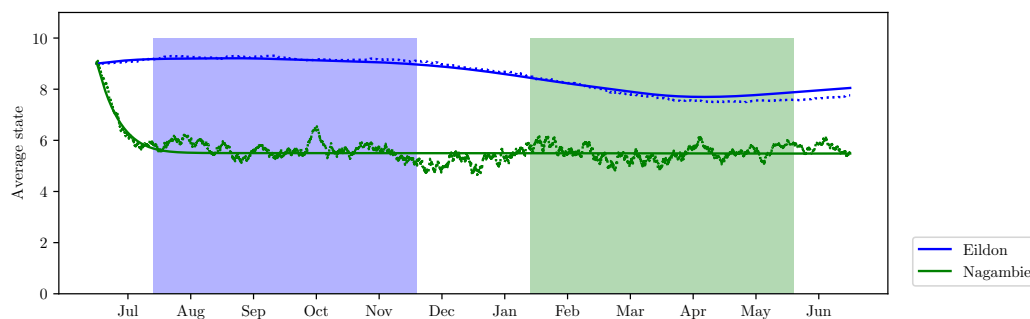

Figure 7. Average levels of Lakes Eildon and Nagambie for program control.

6. Conclusions

The paper presents a framework of controlled coupled MC optimization. The specific character of the problem requires the usage of tensor representation of the set of MCs, which permits us to derive the joint set of DPEs and organize the calculation in a parallel manner. This is important, especially for the most time-consuming minimization of the right-hand side in each discretization step of the numerical solution of the DPE. For a demonstration of the method, we chose a model of the Goulburn River cascade. The model was deliberately simplified to provide a clear demonstration of the approach; however, it is just a first step in a possible detailed analysis of this dam system, which would rely on more detailed data on natural and agricultural processes in the district.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

DPE	Dynamic programming equation
GMW	Goulburn-Murray water
GSM	Goulburn simulation model
MC	Markov chain
ODE	Ordinary differential equation
RHS	Right-hand side

Appendix A. Tensor Product Properties

In this Section we provide the tensor related definitions used throughout the paper. For more details we refer the reader to Reference [27].

For the present paper it is sufficient to define a tensor by its entries like an M-dimensional array of elements $\mathbf{V}^{i^1\dots i^M} \in \mathbb{R}$, where $i^1 \dots i^M$ is multi-index (tuple) given by indices $i^m \in 1, \dots, N^m$. The set of such objects we denote as $\mathbb{R}^{N^1 \times \cdots \times N^m}$.

Definition A1. *Let $V_m \in \mathbb{R}^{N^m}$, $m = 1, \dots, M$ then $\mathbf{V} \in \mathbb{R}^{N^1 \times \cdots \times N^m}$ is called tensor product of vectors V_m*

$$\mathbf{V} = V_1 \otimes \dots \otimes V_M = \bigotimes_{m=1}^{M} V_m$$

if its entries are equal to $\mathbf{V}^{i^1\dots i^M} = V_1^{i^1} \cdot \dots \cdot V_M^{i^M}$.

From Definition A1 it follows that the tensor space $\mathbb{R}^{N^1 \times \cdots \times N^m}$ can be defined as a set of all finite linear combinations of tensor products of vectors from vector spaces $\mathbb{R}^{N^1}, \dots, \mathbb{R}^{N^M}$:

$$\mathbb{R}^{N^1 \times \cdots \times N^m} = \bigotimes_{m=1}^{M} \mathbb{R}^{N^m} = span\{V_1 \otimes \dots \otimes V_M, \text{ where } V_m \in \mathbb{R}^{N^m}, \ m = 1, \dots, M\}.$$

Definition A2. *Let $A_m \in \mathbb{R}^{N^m \times Q^m}$ be a set of matrices, then the tensor product of these matrices is the linear mapping*

$$\mathbf{A} = A_1 \otimes \dots \otimes A_M = \bigotimes_{m=1}^{M} A^m : \ \mathbb{R}^{N^1 \times \cdots \times N^m} \longrightarrow \mathbb{R}^{Q^1 \times \cdots \times Q^m}$$

defined for any $\mathbf{V} \in \mathbb{R}^{N^1 \times \cdots \times N^m}$ by

$$\mathbf{AV} = \left(\bigotimes_{m=1}^{M} A^m\right) \left(\bigotimes_{m=1}^{M} V_m\right) = \bigotimes_{m=1}^{M} (A_m V_m) \in \mathbb{R}^{Q^1 \times \cdots \times Q^m}.$$

References

1. Filar, J.; Vrieze, K. *Competitive Markov Decision Processes*; Springer: New York, NY, USA, 1997. [CrossRef]
2. Williams, B.K. Markov decision processes in natural resources management: Observability and uncertainty. *Ecol. Model.* **2015**, *220*, 830–840. [CrossRef]
3. Delebecque, F.; Quadrat, J.P. Optimal control of markov chains admitting strong and weak interactions. *Automatica* **1981**, *17*, 281–296. [CrossRef]
4. Miller, B.M.; McInnes, D.J. Management of dam systems via optimal price control. *Procedia Comput. Sci.* **2011**, *4*, 1373–1382. [CrossRef]
5. McInnes, D.; Miller, B. Optimal control of time-inhomogeneous Markov chains with application to dam management. In Proceedings of the 2013 Australian Control Conference, Perth, Australia, 4–5 November 2013; pp. 230–237. [CrossRef]
6. Miller, A.; Miller, B. Control of connected Markov chains. Application to congestion avoidance in the Internet. In Proceedings of the 2011 50th IEEE Conference on Decision and Control and European Control Conference, Orlando, FL, USA, 12–15 December 2011; pp. 7242–7248. [CrossRef]
7. Miller, B.M. Optimization of queuing system via stochastic control. *Automatica* **2009**, *45*, 1423–1430. [CrossRef]
8. Miller, A. Using methods of stochastic control to prevent overloads in data transmission networks. *Autom. Remote. Control* **2010**, *71*, 1804–1815. [CrossRef]
9. Miller, B.; McInnes, D. Optimal management of a two dam system via stochastic control: Parallel computing approach. In Proceedings of the 2011 50th IEEE Conference on Decision and Control and European Control Conference, Orlando, FL, USA, 12–15 December 2011; pp. 1417–1423. [CrossRef]
10. Elliott, R.J.; Aggoun, L.; Moore, J.B. *Hidden Markov Models. Estimation and Control*; Stochastic Modelling and Applied Probability; Springer: New York, NY, USA, 1995. [CrossRef]
11. Miller, A.; Miller, B.; Popov, A.; Stepanyan, K. Towards the development of numerical procedure for control of connected Markov chains. In Proceedings of the 2015 5th Australian Control Conference (AUCC), Gold Coast, Australia, 5–6 November 2015; pp. 336–341.
12. Fourneau, J.; Plateau, B.; Stewart, W. Product form for Stochastic Automata Networks. In Proceedings of the 2nd International ICST Conference on Performance Evaluation Methodologies and Tools (Valuetools), Nantes, France, 23–25 October 2007. [CrossRef]
13. Jacod, J.; Shiryaev, A. Limit Theorems for Stochastic Processes. In *Grundlehren Der Mathematischen Wissenschaften*; Springer: Berlin/Heidelberg, Germany, 2003. [CrossRef]
14. Metivier, M.; Pellaumail, J. Stochastic Integration. In *Probability and Mathematical Statistics*; Academic Press: New York, NY, USA, 1980. [CrossRef]
15. Miller, B.; Miller, G.; Siemenikhin, K. Towards the optimal control of Markov chains with constraints. *Automatica* **2010**, *46*, 1495–1502. [CrossRef]
16. Wan, C.B.; Davis, M.H.A. Existence of Optimal Controls for Stochastic Jump Processes. *SIAM J. Control Optim.* **1978**, *17*, 511–524. [CrossRef]
17. Miller, A.; Miller, B. Application of stochastic control to analysis and optimization of TCP. In Proceedings of the 2013 Australian Control Conference, Victoria, Australia, 4–5 November 2013; pp. 238–243. [CrossRef]
18. Schreider, S.; Whetton, P.; Jakeman, A.; Pittock, A. Runoff modelling for snow-affected catchments in the Australian alpine region, eastern Victoria. *J. Hydrol.* **1997**, *200*, 1–23. [CrossRef]
19. Schreider, S.; Jakeman, A.; Letcher, R.; Nathan, R.; Neal, B.; Beavis, S. Detecting changes in streamflow response to changes in non-climatic catchment conditions: Tarm dam development in the Murray–Darling basin, Australia. *J. Hydrol.* **2002**, *262*, 84–98. [CrossRef]
20. Griffith, M.; Codner, G.; Weinmann, E.; Schreider, S. Modelling hydroclimatic uncertainty and short-run irrigator decision making: The Goulburn system. *Aust. J. Agric. Resour. Econ.* **2009**, *53*, 565–584. [CrossRef]
21. Cui, J.; Schreider, S. Modelling of pricing and market impacts for water options. *J. Hydrol.* **2009**, *371*, 31–41. [CrossRef]
22. Dixon, P.; Schreider, S.; Wittwer, G. Combining engineering-based water models with a CGE model. In *Quantitative Tools in Microeconomic Policy Analysis*; Pincus, J., Ed.; Australian Productivity Commission: Canberra, Australia, 2005; Chapter 2, pp. 17–30.

23. Schreider, S.; Weinmann, P.E.; Codner, G.; Malano, H.M. Integrated Modelling System for Sustainable Water Allocation Planning. In Proceedings of the International Congress on Modelling and Simulation MODSIM01, Canberra, Australia, 1 January 2001; Ghassemi, F., McAleer, M., Oxley, L., Scoccimarro, M., Eds.; 2001; Volume 3, pp. 1207–1212.

24. Perera, B.; James, B.; Kularathna, M. Computer software tool REALM for sustainable water allocation and management. *J. Environ. Manag.* **2005**, *77*, 291–300. [CrossRef] [PubMed]

25. Water Measurement Information System of the Department of Environment, Land, Water & Planning of Victoria State Government. Available online: https://data.water.vic.gov.au/ (accessed on 25 July 2020).

26. Climate Data Online Service of the Bureau of Meteorology of the Australian Government. Available online: http://www.bom.gov.au// (accessed on 25 July 2020).

27. Hackbusch, W. *Tensor Spaces and Numerical Tensor Calculus*; Springer Series in Computational Mathematics; Springer: Berlin/Heidelberg, Germany, 2012. [CrossRef]