Sequential Experimental Design for Optimal Structural Intervention in Gene Regulatory Networks Based on the Mean Objective Cost of Uncertainty

Cancer Informatics
Volume 17: 1–10
© The Author(s) 2018
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1176935118790247
SAGE

Mahdi Imani[1], Roozbeh Dehghannasiri[2], Ulisses M Braga-Neto[1,3] and Edward R Dougherty[1,3]

[1]Department of Electrical & Computer Engineering, Texas A&M University, College Station, TX, USA. [2]School of Medicine, Stanford University, Stanford, CA, USA. [3]Center for Bioinformatics and Genomic Systems Engineering, Texas A&M Engineering Experiment Station (TEES), College Station, TX, USA.

ABSTRACT: Scientists are attempting to use models of ever-increasing complexity, especially in medicine, where gene-based diseases such as cancer require better modeling of cell regulation. Complex models suffer from uncertainty and experiments are needed to reduce this uncertainty. Because experiments can be costly and time-consuming, it is desirable to determine experiments providing the most useful information. If a sequence of experiments is to be performed, experimental design is needed to determine the order. A classical approach is to maximally reduce the overall uncertainty in the model, meaning maximal entropy reduction. A recently proposed method takes into account both model uncertainty and the translational objective, for instance, optimal structural intervention in gene regulatory networks, where the aim is to alter the regulatory logic to maximally reduce the long-run likelihood of being in a cancerous state. The mean objective cost of uncertainty (MOCU) quantifies uncertainty based on the degree to which model uncertainty affects the objective. Experimental design involves choosing the experiment that yields the greatest reduction in MOCU. This article introduces finite-horizon dynamic programming for MOCU-based sequential experimental design and compares it with the greedy approach, which selects one experiment at a time without consideration of the full horizon of experiments. A salient aspect of the article is that it demonstrates the advantage of MOCU-based design over the widely used entropy-based design for both greedy and dynamic programming strategies and investigates the effect of model conditions on the comparative performances.

KEYWORDS: entropy, experimental design, dynamic programming, gene regulatory network, greedy search, mean objective cost of uncertainty, structural intervention

RECEIVED: April 3, 2018. **ACCEPTED:** June 25, 2018.

TYPE: Signal Processing Applications In Genomics - Review

FUNDING: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The authors acknowledge the support of the National Science Foundation, through NSF award CCF-1718924.

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: Mahdi Imani, Department of Electrical & Computer Engineering, Texas A&M University, College Station, TX 77843, USA.
Email: m.imani88@tamu.edu

Introduction

A basic problem in genomic signal processing is to derive intervention strategies for gene regulatory networks (GRNs) to avoid undesirable states, in particular, cancerous phenotypes. The problem goes back to the early days of genomics when 2 paradigms were introduced to force dynamical GRNs away from carcinogenic states, *dynamical intervention*[1–3] and *structural intervention*.[4] Substantial work has been done since then (see the work by Dougherty et al[5] for reviews). The goal of dynamical intervention is to find a proper finite or infinite-horizon control strategy for altering the regulatory output of one or more genes at each time point. The goal of structural intervention, which is the focus of this article, is to find a one-time change in regulatory function for beneficially changing the steady-state distribution (SSD) of a GRN. A solution for optimal structure intervention via representation of logical alterations of the regulatory functions is discussed in Xiao et al[6] in the context of probabilistic Boolean networks (PBNs).[7] A solution that applies Markov chain perturbation theory to Markovian regulatory networks to find a structural intervention that optimally reduces the steady-state mass of undesirable states is presented in the work by Qian et al.[8]

The basic theory of structural intervention provides optimal intervention under the assumption that the regulatory model is known; however, in practice, this is generally not the case. For instance, in a Boolean network (BN), or more generally a PBN, it is commonly the case that certain regulatory relations are unknown or at least not known with certainty. In this case, one needs to reformulate the optimization problem to take into account the uncertainty. While we are focusing here on GRNs, this is a problem recognized as far back as the 1960s in control theory.[9–11] More recently, it has been treated in signal processing, first in a minimax framework[12–14] and then in a Bayesian framework,[15,16] and it has also studied in pattern recognition.[17,18] In the case of GRNs, the problem has been addressed in the work by Yoon et al[19] using the mean objective cost of uncertainty (MOCU), which quantifies the uncertainty based on its effect on the objective, in this instance, the degree to which the uncertainty reduces the phenotypical effect of the intervention.

In all cases, one would like to reduce the uncertainty in the model system to achieve better optimal performance. Owing to cost and time, it is prudent to prioritize potential

experiments based on the information they provide and then conduct the most informative. This process is called *experimental design*. Various methods employ entropy,[20–22] the MOCU,[23–25] and the knowledge gradient.[26] This article provides a comparison of entropy-based and MOCU-based methods.

Because uncertainty can be quantified via entropy, a historical approach to experimental design has been to choose an experiment that maximally reduces entropy.[20,21] Assuming that the true model lies in an uncertainty class of models governed by a probability distribution, from a class of potential experiments, the aim is to choose the one that provides model information resulting in the greatest reduction of entropy relative to the distribution. In the case of a GRN, uncertainty might relate to lack of knowledge concerning regulatory relations and the uncertainty class would consist of a collection of GRNs with differing regulatory relations among some of the genes. Potential experiments would characterize unknown regulatory relations, thereby reducing the uncertainty class and lowering entropy.

Entropy does not take into account the objective for building a model. An experiment might reduce entropy but have little or no effect on knowledge necessary for accomplishing the desired objective. In the case of GRNs, structural intervention involves altering gene regulation so as to reduce the steady-state probability of undesirable states, such as cell proliferative (carcinogenic) states. Given a model, one finds an optimal structural intervention.[8] When there is model uncertainty, we desire to reduce uncertainty relevant to determining an optimal structural intervention. The MOCU,[19] which provides a quantification of uncertainty based on the degree to which model uncertainty affects the translational objective, is used for experimental design: choose the experiment that yields the greatest reduction in MOCU.[23]

Typically, one might perform a sequence of experiments to progressively reduce the uncertainty. What should be the order of the experiments to get the best reduction in uncertainty? Using MOCU, this problem has been addressed in the context of GRNs in a greedy sequential fashion[19]: at each step, the optimal experiment is chosen from among the ones not yet performed. In this article, we demonstrate the strength of MOCU-based experimental design by considering optimal sequential experimental design in the context of structural intervention in BN with perturbation (BNp). In this framework, there are M unknown regulatory relations and the aim is to sequentially choose optimal experiments to reduce uncertainty. Notice that "sequential" refers to the step-wise experiments that are taken over the uncertain regulations, which then lead to a better but still one-time structural intervention. We compare design via MOCU and entropy for both greedy and dynamic programming–based sequential design, which has not been previously used with MOCU.

GRNs and Interventions

BNs: a brief overview

Several BN models have been developed in recent years for studying the dynamics of GRNs,[1,27,28] for instance, the cell cycle in the Drosophila fruit fly,[29] in the *Saccharomyces cerevisiae* yeast,[30] and the mammalian cell cycle.[31]

A binary BN on n genes is represented by a set of gene expression, $\{X_1, X_2, \ldots, X_n\}$ and a set of Boolean functions, $\{f_1, \ldots, f_n\}$, that gives functional relationships between the genes over time. The state of each gene is represented by 0 (OFF) or 1 (ON), where $X_i = 1$ and $X_i = 0$ correspond to the activation and inactivation of gene i, respectively. The states of genes at time step k is denoted by a vector $X(k) = (X_1(k), \ldots, X_n(k))$. The value of the ith gene at time step $k+1$ is affected by the value of the k_i predictor genes at time step k via $X_i(k+1) = f_i(X_{i_1}(k), X_{i_2}(k), \ldots, X_{i_{k_i}}(k))$, for $i = 1, \ldots, n$. In a BNp, the state value of each gene at each time point is assumed to be flipped with a small probability p. This produces a dynamical model $X(k+1) = F(X(k)) \oplus \eta(k)$, where $\mathbf{F} = (f_1, \ldots, f_n)$, $\oplus$ is component-wise modulo-2 addition, and $\eta(k) = (\eta_1(k), \eta_2(k), \ldots, \eta_n(k))$ with $\eta_i(k) \sim \text{Bernoulli}(p)$, for $i = 1, \ldots, n$. Letting $\{\mathbf{x}^1, \ldots, \mathbf{x}^{2^n}\}$ be the set of all possible Boolean states, the transition probability matrix (TPM) $\mathbf{P}$ of the Markov chain defined by the state model is given by

$$p_{ij} = P\left(\mathbf{X}(k+1) = \mathbf{x}^i \mid \mathbf{X}(k) = \mathbf{x}^j\right) = p^{||\mathbf{x}^i \oplus \mathbf{F}(\mathbf{x}^j)||_1}(1-p)^{n-||\mathbf{x}^i \oplus \mathbf{F}(\mathbf{x}^j)||_1} \quad (1)$$

for $i, j = 1, \ldots, 2^n$, where p_{ij} refers to the element in the ith row and jth column of the TPM $\mathbf{P}$, and $||\cdot||_1$ is the L_1 norm. For a nonzero perturbation process $(p > 0)$, the corresponding Markov chain of a BNp possesses a SSD π describing the long-run behavior of system. The SSD can be computed based on the TPM of the Markov chain as $\pi^T = \pi^T \mathbf{P}$, where $\mathbf{v}^T$ denotes the transpose of $\mathbf{v}$ and the ith element denotes the steady-state probability of being at state $\mathbf{x}^i$.

Structural intervention of GRNs

We now briefly review the solution that applies Markov chain perturbation theory to the TPM to find a structural intervention that optimally reduces the steady-state mass of undesirable states.[8] Given a known BNp, under the rank 1 function perturbation, the TPM $\mathbf{P}$ will be altered to $\tilde{\mathbf{P}} = \mathbf{P} + \mathbf{a}\mathbf{b}^T$,[8] where $\mathbf{a}$ and $\mathbf{b}$ are arbitrary vectors and $\mathbf{b}^T\mathbf{e} = 0$ ($\mathbf{e}$ is the all unity column vector). A special case of a rank 1 perturbation, called a single-gene perturbation process, is considered in this article. According to this process, the output state of a single-input state changes and the output states of other states stay unchanged. Let Ψ be a class of potential interventions to the network. Let $\tilde{\mathbf{F}} = (\tilde{f}_1, \ldots, \tilde{f}_n)$ be the Boolean function after intervention. A

single-gene perturbation for the input state j changes the output value of the Boolean function: $\mathbf{x}^s = \tilde{\mathbf{F}}(\mathbf{x}^j) \neq \mathbf{F}(\mathbf{x}^j) = \mathbf{x}^r$, and $\tilde{\mathbf{F}}(\mathbf{x}^i) = \mathbf{F}(\mathbf{x}^i)$, for $i = 1,\ldots,2^n$ and $i \neq j$. We shall refer to this as a (j,s) intervention. The TPM after perturbation, $\tilde{\mathbf{P}}$, is the same as $\mathbf{P}$, except for $\tilde{p}_{jr} = p_{jr} - (1-p)^n$ and $\tilde{p}_{js} = p_{js} + (1-p)^n$. The SSD of the system after perturbation can be computed as follows[8]

$$\tilde{\pi}_i(j,s) = \pi_i + \frac{(1-p)^n \pi_j (z_{si} - z_{ri})}{1-(1-p)^n (z_{sj} - z_{rj})} \tag{2}$$

where $\tilde{\pi}_i(j,s)$ is the steady-state probability of the ith state of the perturbed system following a (j,s) intervention, $\mathbf{Z} = [\mathbf{I} - \mathbf{P} + \mathbf{e}\pi^T]^{-1}$ is the fundamental matrix of a BNp, $\mathbf{I}$ being the $n \times n$ identity matrix, and $z_{si}, z_{ri}, z_{sj}, z_{rj}$ are elements of $\mathbf{Z}$.

If U is the set of undesirable Boolean states, then $\tilde{\pi}_U(j,s) = \sum_{i \in U} \tilde{\pi}_i(j,s)$ is the steady-state probability mass of undesirable states after applying a (j,s) intervention. The optimal single-gene perturbation structural intervention (j^*, s^*) minimizes $\tilde{\pi}_U(j,s)$:

$$(j^*, s^*) = \underset{j,s \in \{1,2,\ldots,2^n\}}{\operatorname{argmin}} \tilde{\pi}_U(j,s) \tag{3}$$

Experimental Design

The complex regulatory machinery of the cell and the lack of sufficient data for accurate inference create significant uncertainty in GRN models. Consider a GRN possessing M uncertain parameters $\theta^1, \theta^2, \ldots, \theta^M$. In our application, θ^i corresponds to a regulatory relation of an uncertain type that can take on 2 different values: "$\mathcal{A}$" for activating regulation and "$\mathcal{S}$" for suppressive regulation. These unknown parameters result in 2^M different BN models for the system that differs in one or more of these uncertain regulations. Let $\Theta = \{\theta_1, \ldots, \theta_{2^M}\}$ be the uncertainty class of these network models, where $\theta_j \in \{\mathcal{A}, \mathcal{S}\}^M$, for $j = 1,\ldots,2^M$. The prior distribution over BN models can be encoded into a single column vector:

$$p(0) = \left[P(\theta^* = \theta_1), \ldots, P(\theta^* = \theta_{2^M}) \right]^T$$

where θ^* is a vector containing the true values of the parameters.

For a given initial distribution $p(0)$ and $i = 1,\ldots,M$, the prior probability that the ith regulation is activating is

$$P(\theta^i = \mathcal{A}) = E_{p(0)}\left[1_{\theta(i)=\mathcal{A}}\right] = \sum_{j=1}^{2^M} p_j(0) 1_{\theta_j(i)=\mathcal{A}}$$

where $1_{\theta_j(i)=\mathcal{A}} = 1$ if $\theta_j(i) = \mathcal{A}$ and 0 otherwise. The *initial belief state* is $b(0) = [P(\theta^1 = \mathcal{A}), \ldots, P(\theta^M = \mathcal{A})]^T$.

As in the work by Dehghannasiri et al,[23] we assume that there exist M experiments $T_1, \ldots, T_M$, where T_i determines the regulation θ^i. More general experimental formulations are possible; for instance, there is a probability that T_i can return the wrong value.[25]

Let $\mathbf{b}(k)$ be the belief state before conducting the kth experiment. Given that experiment T_i at time step k is performed, if the outcome of the experiment shows that $\theta^i = \mathcal{A}$, then the ith element of the belief vector at time step $k+1$ will get the value 1 and the other elements will get their previous values: $\mathbf{b}_i(k+1) = 1$, $\mathbf{b}_l(k+1) = \mathbf{b}_l(k)$ for $l = 1,\ldots,M, l \neq i$. However, if $\theta^i = \mathcal{S}$, then the ith element of the belief vector will be 0 and the rest will be unaltered from time k to $k+1$.

Thus, each of the M elements of the belief vector can take 3 possible values during the experimental design process and the belief vector is of the form $\mathbb{B} = [\mathbf{b}^1, \ldots, \mathbf{b}^{3^M}]$, where $\mathbf{b}(k) \in \mathbb{B}$. We view transition in the belief space as a Markov decision process with 3^M states. The *controlled transition matrix* in the belief space under experiment T_i is a matrix of size $3^M \times 3^M$. The element associated with the probability of transition from state $\mathbf{b} \in \mathbb{B}$ to state $\mathbf{b}' \in \mathbb{B}$ under experiment T_i can be written as follows:

$$\begin{aligned} \mathbf{Tr}_{bb'}(T_i) &= P\left(\mathbf{b}(k+1) = \mathbf{b}' \,|\, \mathbf{b}(k) = \mathbf{b}, T_i\right) \\ &= \begin{cases} \mathbf{b}_i & \text{If } \mathbf{b}'_i = 1 \text{ and } \mathbf{b}_l = \mathbf{b}'_l \text{ for } l \neq i \\ 1 - \mathbf{b}_i & \text{If } \mathbf{b}'_i = 0 \text{ and } \mathbf{b}_l = \mathbf{b}'_l \text{ for } l \neq i \\ 0 & \text{o.w.} \end{cases} \end{aligned} \tag{4}$$

Greedy-MOCU

Optimal experimental design using the MOCU, first proposed in the work by Dehghannasiri et al,[23] is briefly described in this section. Let $\xi_\theta(\psi)$ be the cost of applying the intervention $\psi \in \Psi$ to the network $\theta \in \Theta$. Using equation (3), for any $\theta \in \Theta$, the optimal single-gene perturbation structural intervention for a BNp defined by a given uncertainty vector θ is $\psi_\theta = (j^*_\theta, s^*_\theta)$, where $\xi_\theta(\psi) \geq \xi_\theta(\psi_\theta)$ for any $\psi \in \Psi$.

The MOCU relative to an uncertainty class represented by the belief vector $\mathbf{b}$ and a class Ψ of interventions is defined as follows:

$$\begin{aligned} M_\Psi(\Theta \,|\, \mathbf{b}) &= E_{\Theta|\mathbf{b}}\left[\xi_\theta(\psi^{\theta|\mathbf{b}}_{\mathrm{IBR}}) - \xi_\theta(\psi_\theta)\right] \\ &= \sum_{j=1}^{2^M} p^{\mathbf{b}}_j \left[\xi_{\theta_j}(\psi^{\theta|\mathbf{b}}_{\mathrm{IBR}}) - \xi_{\theta_j}(\psi_{\theta_j})\right] \end{aligned} \tag{5}$$

where $\psi^{\theta|\mathbf{b}}_{\mathrm{IBR}}$ is an *intrinsically Bayesian robust* (IBR) intervention,

$$\psi^{\theta|\mathbf{b}}_{\mathrm{IBR}} = \underset{\psi \in \Psi}{\operatorname{argmin}} E_{\mathbf{b}}[\xi_\theta(\psi)] = \underset{\psi \in \Psi}{\operatorname{argmin}} \sum_{j=1}^{2^M} p^{\mathbf{b}}_j \xi_{\theta_j}(\psi) \tag{6}$$

and $p^{\mathbf{b}}$ is the vector of posterior probabilities of network models for a belief vector $\mathbf{b}$, which can be computed based on the independency of the regulations as follows:

$$p_j^{\mathbf{b}} = \prod_{i=1}^{M}\left[\mathbf{b}_i 1_{\theta_j(i)=\mathcal{A}} + (1-\mathbf{b}_i)1_{\theta_j(i)=\mathcal{S}}\right] \quad (7)$$

for $j = 1,\ldots,2^M$. The IBR intervention $\psi_{\mathrm{IBR}}^{\theta|\mathbf{b}}$ depends on the belief state $\mathbf{b}$, whereas the optimal intervention ψ_θ is designed for a specific network model $\theta \in \Theta$. The MOCU is the expected cost increase that results from applying a robust intervention over all networks in θ instead of the optimal intervention for the unknown true network.

The goal of sequential greedy MOCU-based experimental design[23]), referred to herein as Greedy-MOCU, is to select an experiment at each time point that results in the maximal reduction in MOCU in the next time step. If $\mathbf{b}$ is the current belief state, then the Greedy-MOCU decision is given by

$$\begin{aligned} i^* &= \underset{i\in\{1,\ldots,M\}}{\operatorname{argmin}} E_{\mathbf{b}'|\mathbf{b},T_i}\left[M_\Psi(\Theta\,|\,\mathbf{b}') - M_\Psi(\Theta\,|\,\mathbf{b})\right] \\ &= \underset{i\in\{1,\ldots,M\}}{\operatorname{argmin}} \sum_{\mathbf{b}'\in\mathbb{B}} \mathbf{Tr}_{\mathbf{b}\mathbf{b}'}(T_i) M_\Psi(\Theta\,|\,\mathbf{b}') \end{aligned} \quad (8)$$

where the second equality follows by expressing the expectation $E_{\mathbf{b}'|\mathbf{b},T_i}$ in terms of $\mathbf{Tr}_{\mathbf{b}\mathbf{b}'}(T_i)$, for $\mathbf{b}' \in \mathbb{B}$, and then dropping the terms unrelated to minimization. After making the decision and observing outcomes, one needs to update the belief state and repeat another experimental design process if necessary.

Dynamic programming MOCU

Greedy-MOCU uses the expected value of MOCU in the next time step for decision making at the current time. If the number of experiments, N, is known a priori, then all future experiments (the remaining horizon) can be taken into account during the decision-making process. In this section, we introduce optimal finite-horizon experimental design based on dynamic programming (DP),[32] which we call DP-MOCU.

Let $\mu_k(\mathbf{b})$ be a policy at time step k that maps a belief vector $\mathbf{b} \in \mathbb{B}$ into an experiment in $\{T_1,\ldots,T_M\}$. We define a bounded *immediate cost function* at time step k corresponding to transition from the belief vector $\mathbf{b}(k) = \mathbf{b}$ into the belief vector $\mathbf{b}(k+1) = \mathbf{b}'$ under policy μ_k as follows:

$$g_k\left(\mathbf{b},\mathbf{b}',\mu_k(\mathbf{b})\right) = M_\Psi(\Theta\,|\,\mathbf{b}') - M_\Psi(\Theta\,|\,\mathbf{b})$$

for $k = 0,\ldots,N-1$, where $g_k(\mathbf{b},\mathbf{b}',\mu_k(\mathbf{b})) \le 0$. The *terminal cost function* is defined as $g_N(\mathbf{b}) = M_\Psi(\Theta\,|\,\mathbf{b})$, for any $\mathbf{b} \in \mathbb{B}$.

Letting Π be the space of all possible policies, using the definitions of the immediate and terminal cost functions, an optimal policy, $\mu_{0:N-1}^{\mathrm{MOCU}}$, is given by solving the minimization problem:

$$\underset{\mu_{0:N-1}\in\Pi}{\operatorname{argmin}} E\left[\sum_{k=0}^{N-1} g_k\left(\mathbf{b}(k),\mathbf{b}(k+1),\mu_k(\mathbf{b}(k))\right) + g_N(\mathbf{b}(N))\right] \quad (9)$$

where the expectation is taken over stochasticities in belief transition.

Dynamic programming provides a solution for the minimization in equation (9). The method starts by setting the terminal cost function as $J_N^{\mathrm{MOCU}}(\mathbf{b}) = g_N(\mathbf{b})$ for $\mathbf{b} \in \mathbb{B}$. Then, in a recursively backward fashion, the optimal cost function can be computed as follows:

$$\begin{aligned} J_k^{\mathrm{MOCU}}(\mathbf{b}) &= \min_{i\in\{1,\ldots,M\}} E_{\mathbf{b}'|\mathbf{b},T_i}\left[g_k\left(\mathbf{b},\mathbf{b}',T_i\right) + J_{k+1}^{\mathrm{MOCU}}(\mathbf{b}')\right] \\ &= \min_{i\in\{1,\ldots,M\}}\left[\sum_{\mathbf{b}'\in\mathbb{B}} \mathbf{Tr}_{\mathbf{b}\mathbf{b}'}(T_i)\left(g_k\left(\mathbf{b},\mathbf{b}',T_i\right) + J_{k+1}^{\mathrm{MOCU}}(\mathbf{b}')\right)\right] \end{aligned} \quad (10)$$

with an optimal policy, $\mu_k^{\mathrm{MOCU}}(\mathbf{b})$, given by

$$\underset{i\in\{1,\ldots,M\}}{\operatorname{argmin}}\left[\sum_{\mathbf{b}'\in\mathbb{B}} \mathbf{Tr}_{\mathbf{b}\mathbf{b}'}(T_i)\left(g_k\left(\mathbf{b},\mathbf{b}',T_i\right) + J_{k+1}^{\mathrm{MOCU}}(\mathbf{b}')\right)\right] \quad (11)$$

for $\mathbf{b} \in \mathbb{B}$ and $k = N-1,\ldots,0$, where $\mathbf{Tr}(T_i)$ is defined in equation (4). Unlike Greedy-MOCU, the DP-MOCU policy decides which uncertain regulation should be determined at each step to maximally reduce the uncertainty relative to the objective after conducting all N experiments.

Greedy entropy

The idea of entropy-based experimental design[20,21] is to reduce the amount of the entropy, which quantifies the uncertainty of the system. While MOCU-based techniques take action to reduce the uncertainty with respect to an objective, entropy-based techniques do not take into account the objective during decision making. Performance comparisons are made in section "Results."

The entropy for belief vector $\mathbf{b}$ is $H(\mathbf{b}) = -\sum_{j=1}^{2^M} p_j^{\mathbf{b}} \log_2 p_j^{\mathbf{b}}$, where $p^{\mathbf{b}}$ is the posterior probability of network models under the belief state $\mathbf{b}$ defined in equation (7). The maximum value of the entropy is M, which corresponds to a uniform prior over the network models, and the minimum value is 0, which corresponds to certainty.

The Greedy-Entropy approach sequentially chooses an experiment to minimize the expected entropy at the next time step:

$$\begin{aligned} i^* &= \underset{i\in\{1,\ldots,M\}}{\operatorname{argmin}} E_{\mathbf{b}'|\mathbf{b},T_i}\left[H(\mathbf{b}') - H(\mathbf{b})\right] \\ &= \underset{i\in\{1,\ldots,M\}}{\operatorname{argmin}}\left[-\sum_{\mathbf{b}'\in\mathbb{B}} \mathbf{Tr}_{\mathbf{b}\mathbf{b}'}(T_i)\sum_{j=1}^{2^M} p_j^{\mathbf{b}'}\log_2 p_j^{\mathbf{b}'}\right] \end{aligned} \quad (12)$$

where the second equality is obtained by removing constant terms.

Dynamic programming entropy

The Greedy-Entropy approach takes into account only the entropy in the next step for selecting the experiment to be performed at the current step. If the number N of experiments is known a priori, then the DP technique is used for finding an optimal entropy-based solution. In the work by Huan and Marzouk,[22] an approximate DP solution based on the entropy scheme for cases with a continuous belief space is provided. Here, we employ the optimal DP solution because the belief space is finite. Again letting $\mu_k(\mathbf{b}):\mathbf{b}\to\{T_1,\ldots,T_M\}$ be a policy at time step k, we define a bounded immediate cost function at time step k corresponding to transition from the belief vector $\mathbf{b}(k)=\mathbf{b}$ to the belief vector $\mathbf{b}(k+1)=\mathbf{b}'$ under policy μ_k by

$$\tilde{g}_k(\mathbf{b},\mathbf{b}',\mu_k(\mathbf{b}))=H(\mathbf{b}')-H(\mathbf{b})$$

for $k=0,\ldots,N-1$. Define the terminal cost function by $\tilde{g}_N(\mathbf{b})=H(\mathbf{b})$, for any $\mathbf{b}\in\mathbb{B}$. Using the immediate and terminal cost functions $\tilde{g}_k$ and $\tilde{g}_N$ instead of g_k and g_N, for $k=0,\ldots,N-1$, in the DP process, the optimal finite-horizon policy, $\mu_{0:N-1}^{\text{Entropy}}(\mathbf{b})$, for $\mathbf{b}\in\mathbb{B}$, based on the entropy, is obtained. It is called the DP-Entropy policy.

Results

Simulation setup

According to the majority vote rule for generating BN models of GRNs, the ith Boolean predictor is given by

$$X_i(k+1)=f_i\big(X(k)\big)=\begin{cases}1 & \text{If } \sum_j R_{ij}\,X_j(k)>0\\ 0 & \text{If } \sum_j R_{ij}\,X_j(k)<0\\ X_i(k) & \text{If } \sum_j R_{ij}\,X_j(k)=0\end{cases}$$

for $i=1,\ldots,n$, where R_{ij} can take 3 values: $R_{ij}=+1$ if there is an activating regulation ($\mathcal{A}$) from gene j to gene i, $R_{ij}=-1$ if there is suppressive regulation ($\mathcal{S}$) from gene j to gene i, and $R_{ij}=0$ if gene j is not an input to gene i.[33]

We employ the symmetric Dirichlet distribution for generating the initial distribution over various network models:

$$p^{\mathrm{b}}(0)\sim f(p^{\mathrm{b}}(0);\phi)=\frac{\Gamma\left(\phi 2^M\right)}{\Gamma(\phi)^{2^M}}\prod_{j=1}^{2^M}p_j^{\mathrm{b}}(0)^{\phi-1}$$

where Γ is the gamma function and $\phi>0$ is the parameter of the symmetric Dirichlet distribution. The expected value of the initial distribution for any value of ϕ is a vector of size 2^M with all elements $1/2^M$. ϕ specifies the variability of the initial distributions; the smaller ϕ is, the more the initial distributions deviate from the uniform distribution.

Performance evaluation based on synthetic BNps

To evaluate performance, simulations based on synthetic BNps have been performed. A total of 100 random BNps of size 6 with a single set of M unknown regulations for each network have been considered. The perturbation probability is set to $P=0.001$. Different values of P have been tried and similar results, as presented in the sequel, have been observed. The states with upregulated first gene are assumed to be undesirable ($U=\{\mathbf{x}^1,\ldots,\mathbf{x}^{32}\}$). Three different values are considered for the Dirichlet parameter: $\phi=0.1,1,$ and 10. From each ϕ, 500 initial distributions are generated.

Five experimental design strategies are considered: (1) Greedy-MOCU, (2) DP-MOCU, (3) Greedy-Entropy, (4) DP-Entropy, and (5) Random. A successful experimental design strategy has the ability to effectively reduce the cost of intervention. Thus, the robust intervention based on the resulting belief state of each strategy is applied to the true (unknown) network and the cost of intervention (total steady-state mass in undesirable states) is used as a metric. For a given belief state $\mathbf{b}(k)$ computed before taking the kth experiment, the cost is $\xi_{\theta^*}(\psi_{\mathrm{IBR}}^{\theta|\mathbf{b}(k)})$. $H(\mathbf{b}(k))$ represents the amount of remaining uncertainty in the system for a given belief state $\mathbf{b}(k)$. Thus, we define the intervention gain of conducting the chosen experiment over a random experiment by $\xi_{\theta^*}(\psi_{\mathrm{IBR}}^{\theta|\mathbf{b}^{\mathrm{rnd}}(k)})-\xi_{\theta^*}(\psi_{\mathrm{IBR}}^{\theta|\mathbf{b}(k)})$, and the entropy gain as $H(\mathrm{b}^{\mathrm{rnd}}(k))-H(\mathrm{b}(k))$, where $\mathbf{b}(k)$ is the belief state after performing the kth experiment determined via experimental design (Greedy-MOCU, DP-MOCU, Greedy-Entropy, or DP-Entropy), and $\mathbf{b}^{\mathrm{rnd}}(k)$ is the belief vector before performing the kth experiment during the random experimental design process.

Figure 1 shows the average gain of intervention with respect to the horizon length N strategies for different numbers of unknown regulations (M) and Dirichlet parameters (ϕ). The figure shows curves for $M=2,3,5,7$ and $\phi=0.1,1,10$. The curves end at gain value 0 when $M=N$, so that all regulations have been identified. In practice, the number of unknown regulations is usually more than the number of experiments which can be performed. In these cases, large intervention gains have been attained by the MOCU-based strategies in comparison with entropy-based techniques, thus demonstrating the effectiveness of MOCU-based strategies in reducing network uncertainty relevant to the objective.

The maximum amount of gain in MOCU-based strategies is achieved for $\phi=10$ (maximum uncertainty). In contrast, the intervention gains in entropy-based strategies are very close to 0 when the initial distribution is closer to uniform $(\phi=10)$ and increase as this distribution deviates from uniform $(\phi=0.1)$. Indeed, as ϕ gets larger (initial distributions get closer to uniform), the Entropy scheme does not discriminate between potential experiments and performs like a random selection approach. Thus, entropy-based strategies slightly perform better than the random strategy for nonuniform prior

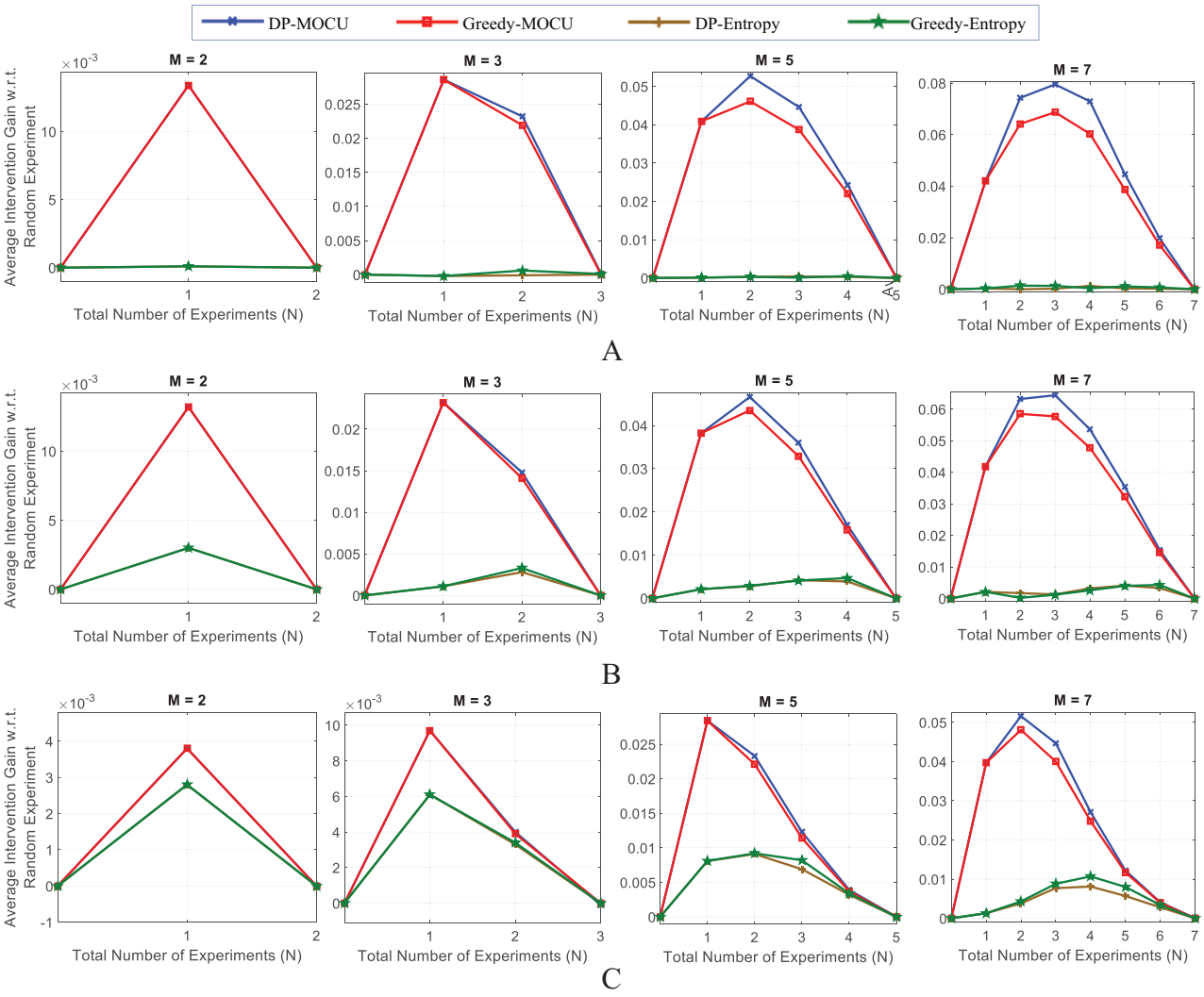

Figure 1. The average intervention gain with respect to the total number of experiments, N, for randomly generated synthetic 6-gene Boolean networks with 2, 3, 5, 7 uncertain regulations (M). (A) $\phi=10$, (B) $\phi=10$, and (C) $\phi=10$.

distributions, with their performance being far worse than MOCU-based techniques. In addition, the peak in the intervention gain is shifted slightly into the left side as ϕ decreases. This is due to the fact that in the presence of a nonuniform initial distribution, the MOCU-based strategies are capable of selecting the first most effective experiments in early steps to reduce the intervention cost.

In Figure 1, DP-MOCU outperforms Greedy-MOCU in all cases with respect to the cost of intervention because future experiments are taken into account for decision making in DP-MOCU, as opposed to Greedy-MOCU, which only considers one-step look-ahead. When the total number of experiments (N) is 1, Greedy-MOCU and DP-MOCU are equivalent and the same gain can be seen for both strategies. The highest gain difference is achieved when the horizon length N is less than M, the number of uncertain parameters.

To better appreciate the performance of entropy-based techniques, the gain in entropy is reported in Figure 2. Once again, the figure shows curves for $M=2,3,5,7$ and $\phi=0.1,1,10$. Comparison between Figures 1 and 2 shows that the maximum entropy reduction by the entropy-based strategies does not necessarily result in the highest reduction in the cost of intervention, which is the main objective of performing experimental design. Interestingly, although DP-Entropy is more successful in reducing the entropy value in comparison with Greedy-Entropy (as expected), it does not outperform Greedy-Entropy relative to average gain in intervention.

Next, we consider the effect of the initial distribution on the performance of various experimental design strategies. The horizon length and the number of unknown regulations are set to be $N=4$ and $M=7$, respectively. The initial distribution is a vector of size 2^7. The entropy of this initial distribution specifies the amount of initial uncertainty in the system. The closer this value is to its maximum value 7, the closer the initial distribution is to the uniform distribution. In the figure, we observe that as the entropy of the initial distribution increases, the performance of both Greedy-MOCU and DP-MOCU increases as well. This growth is higher for DP-MOCU compared

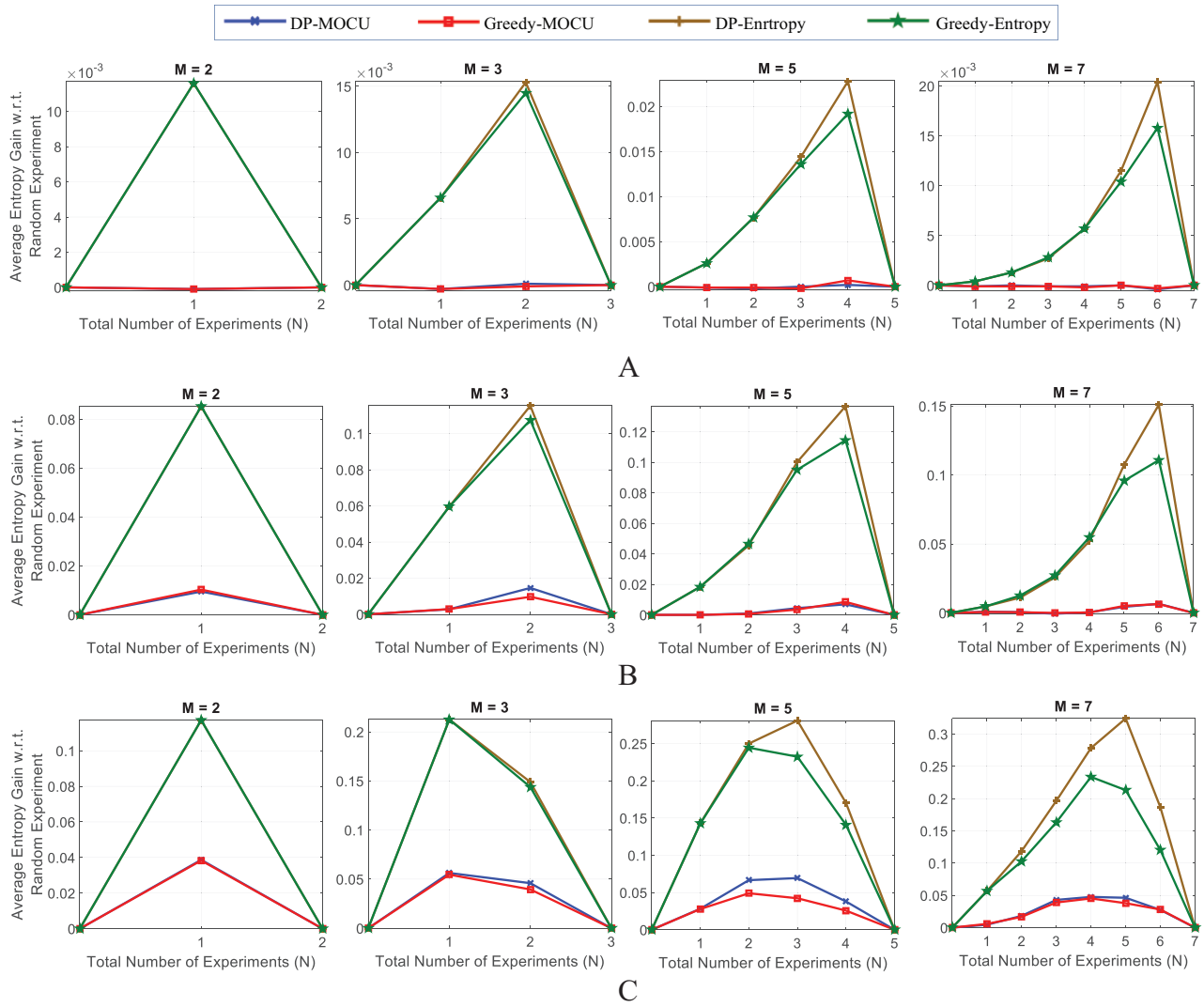

Figure 2. The average entropy gain with respect to the total number of experiments, N , for randomly generated synthetic 6-gene Boolean networks with 2, 3, 5, 7 uncertain regulations (M) . (A) $\phi=10$, (B) $\phi=10$, and (C) $\phi=10$.

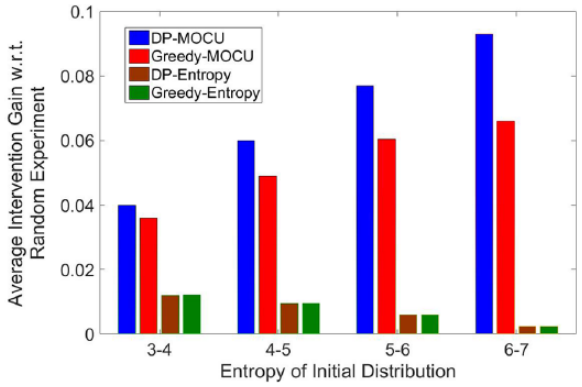

Figure 3. The average intervention gain with respect to the entropy of the initial distribution, $p(0)$, for randomly generated synthetic 6-gene Boolean networks with 7 unknown regulations and total number of experiments (N) equal to 4 .

with Greedy-MOCU, which shows the superiority of DP-MOCU in reducing the intervention cost in the presence of high uncertainty in the system. However, note the reduction trend in the amount of intervention gain for the entropy-based techniques as the entropy of the initial distribution increases. This is due to the fact that the entropy-based strategies are unable to discriminate between potential experiments in the presence of the uniform initial distribution and perform like random selection (Figure 3).

The average cost of robust intervention with respect to the number of conducted experiments for different experimental design strategies is shown in Figure 4. DP-MOCU has the lowest average cost of robust intervention at the end of the horizon (after taking all N experiments); however, Greedy-MOCU has the lowest cost before reaching the end of the horizon. This observation can be understood by looking at the finite-horizon DP policy. DP-MOCU finds a sequence of experiments from time 0 to $N-1$ to minimize the expected sum of the differences of MOCUs throughout this interval. The expected value of MOCU after conducting the last experiment plays the key role in the decision making by the DP

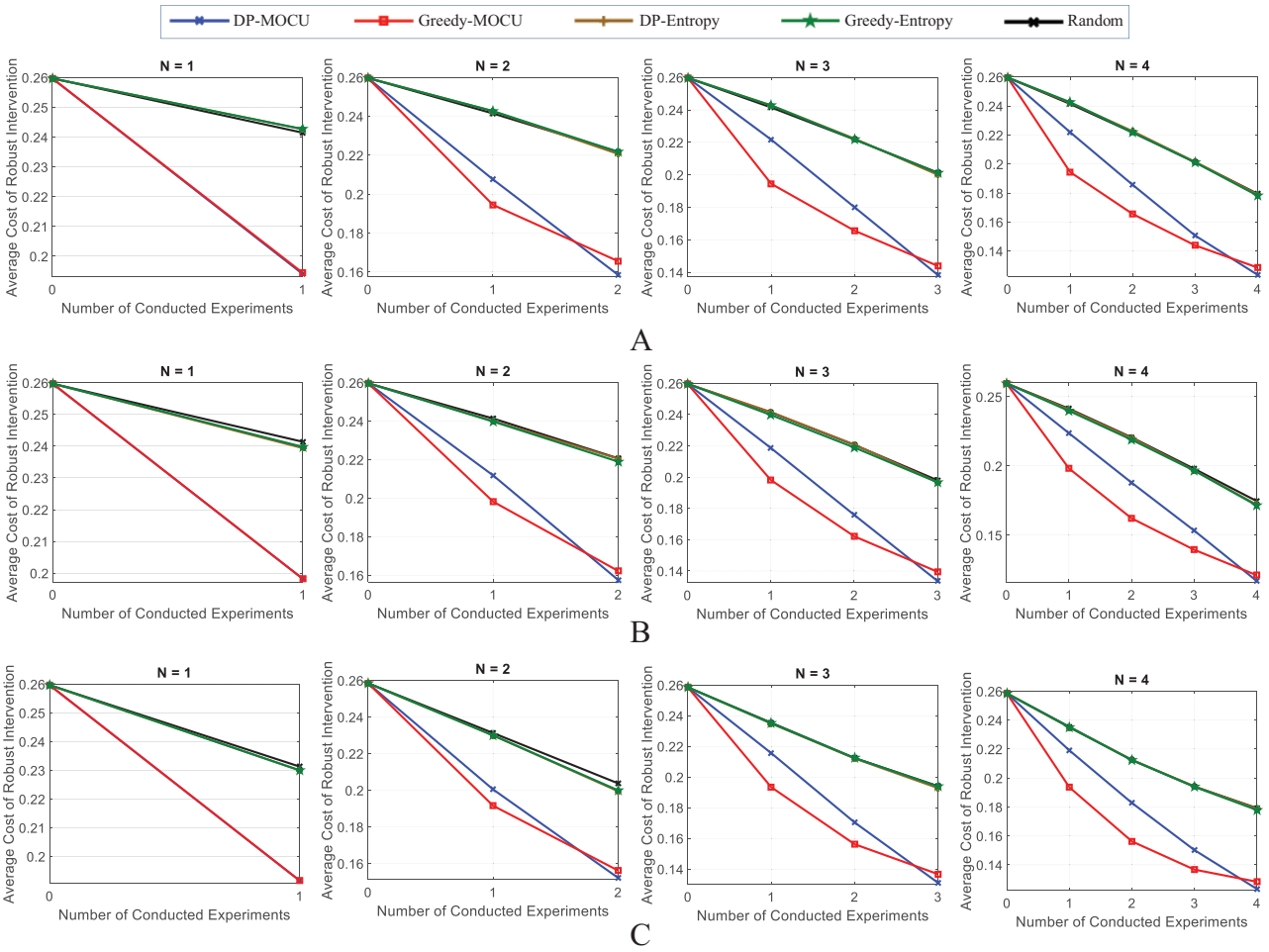

Figure 4. The average cost of robust intervention with respect to the number of conducted experiments obtained by various experimental design strategies for randomly generated synthetic BNps with 7 unknown regulations (M) and $\phi = 0.1,1,10$. (A) $\phi = 10$, (B) $\phi = 10$, and (C) $\phi = 10$.

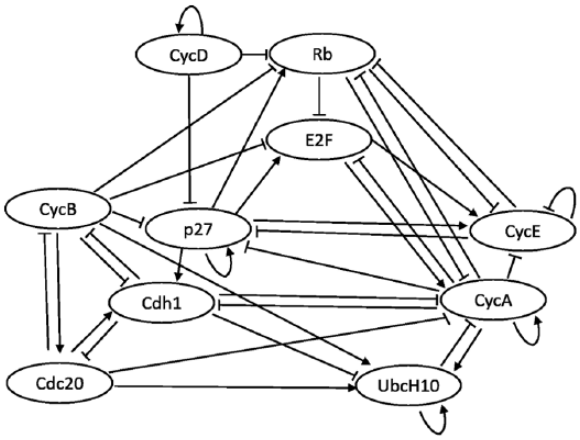

Figure 5. A gene regulatory network model of the mammalian cell cycle. Normal arrows represent activating regulations and blunt arrows represent suppressive regulations.

policy. Thus, the capability of DP-MOCU in planning for reducing MOCU at the end of the horizon, as opposed to Greedy-MOCU which takes only the next step into account for decision making, results in the lowest average cost of robust intervention by DP-MOCU at the end of horizon. Both DP-MOCU and Greedy-MOCU are equivalent for horizon length $N = 1$ and behave differently for other cases. When the number of experiments is not known a priori, Greedy-MOCU may be preferred to DP-MOCU because, as presented in Figure 4, the intervention gain in DP-MOCU might be lower than Greedy-MOCU before conducting the total number of experiments.

Performance evaluation based on the mammalian cell cycle network

The mammalian cell cycle involves a sequence of events resulting in the duplication and division of the cell. It occurs in response to growth factors and under normal conditions; it is a tightly controlled process. A regulatory model for the mammalian cell cycle, proposed in Fauré et al,[31] is shown in Figure 5. This model contains 10 genes: CycD, Rb, p27, E2F, CycE, CycA, Cdc20, Cdh1, UbcH10, and CycB. The blunt and normal arrows represent suppressive $(\mathcal{S})$ and activating $(\mathcal{A})$ regulations, respectively. Mammalian cell division is

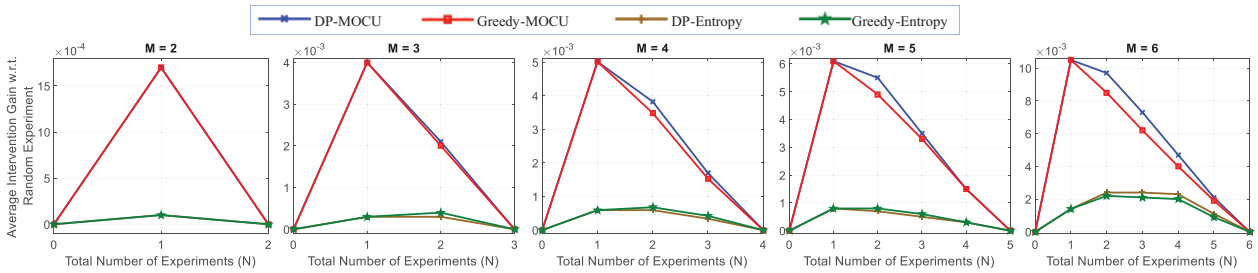

Figure 6. The average intervention gains of various experimental strategies versus the random strategy for the mammalian cell cycle network for 2 to 6 unknown regulations (M) and $\phi = 1$.

coordinated with the overall growth of the organism via extracellular signals that control the activation of CycD in the cell. Cell division happens due to the positive stimuli activating cyclin D (CycD). When CycD is upregulated, it inactivates the tumor suppressor Rb protein via phosphorylation. Rb can also be expressed if gene p27 and either CycE or CycA is active. The activation of Rb in the absence of stimuli causes cell proliferative (cancerous) phenotypes. States with downregulated CycD, Rb, and p27 $(X_1 = X_2 = X_3 = 0)$ are undesirable, representing cancerous phenotypes. The goal is to reduce the steady-state probability mass of the set of undesirable states, $U = \{\mathbf{x}^1, \ldots, \mathbf{x}^{128}\}$, via structural intervention.

We consider various cases with 2 to 6 unknown regulations (M). We randomly select 100 different sets of M regulations from the network, for which we assume their regulatory information is not known and apply various experimental design strategies to predict the experiment to be performed. A total of 500 initial distributions have been generated from the Dirichlet distribution with parameter $\phi = 1$.

The average intervention gains for various experimental design strategies are presented in Figure 6, which shows that curves for $M = 2, \ldots, 6$. DP-MOCU and Greedy-MOCU have the highest average intervention gain in comparison with the entropy-based strategies. DP-MOCU is clearly superior to Greedy-MOCU for cases with larger numbers of unknown regulations and when the number of experiments is smaller than the number of unknown regulations $(1 < N < M)$. Both Greedy-Entropy and DP-Entropy perform poorly in all cases.

Computational complexity analysis

Consider a network with n genes, in which the states 2^{n-1} to 2^n are undesirable. Structural intervention requires 2^{n-1} searches over $2^n \times 2^n$ state pairs. This gives complexity $O(2^{3n})$ for the optimal intervention process for a single network. Given M uncertain parameters, which poses 2^M different network models, the complexity of Greedy-MOCU is of order $O(2^M \times 2^{3n})$. However, DP-MOCU has an extra step for the DP process. The complexity of the DP process is of order $O(3^{2M} \times N)$, where N is the horizon length. Thus, the complexity of DP-MOCU is $O(\max\{3^{2M} \times N, 2^M \times 2^{3n}\})$. In contrast to the MOCU-based strategies, the complexities of the entropy-based techniques are independent of the intervention process. Greedy-Entropy and DP-Entropy have complexities $O(2^M \times 2^n)$ and $O(\max\{3^{2M} \times N, 2^M \times 2^n\})$, respectively.

Table 1. Comparing the approximate processing times (in seconds) of various experimental design methods for networks of size n with M uncertain regulations, and $N = 3$.

		$n = 10$	$n = 11$	$n = 12$
Greedy-MOCU	$M = 4$	250	2651	32933
	$M = 5$	493	5210	65208
	$M = 6$	967	10264	127397
DP-MOCU	$M = 4$	272	2696	32989
	$M = 5$	490	5294	65323
	$M = 6$	1002	10314	127413
DP-Entropy	$M = 4$	5	11	21
	$M = 5$	15	29	63
	$M = 6$	44	86	173
Greedy-Entropy	$M = 4$	6	13	25
	$M = 5$	18	36	69
	$M = 6$	50	99	184

Abbreviations: DP, dynamic programming; MOCU, mean objective cost of uncertainty.

Table 1 shows approximate processing times for networks of different size with various numbers of regulations. Simulations have been run on a machine with 16 GB of RAM and Intel Core i7 CPU, 3.6 GHz. The running time of the MOCU-based strategies grows exponentially as the number of genes increases. It also increases with increases in the number of unknown regulations. It can be seen that the running time of DP-MOCU is slightly higher than that of Greedy-MOCU owing to the extra DP recursion in DP-MOCU.

Clearly, computational complexity is an issue. The issue has been addressed in the context of structural intervention in the work by Dehghannasiri et al,[23] where computation reduction for MOCU-based design is achieved via network reduction schemes. These result in suboptimal experimental design, but they are still superior to random design.

Conclusions

By taking into account the operational objective, MOCU-based experimental design significantly outperforms entropy-based design. Our aim in this article has been 2-fold: to demonstrate this advantage and to propose and examine the effect of using finite-horizon DP for sequential design. The simulations show that if one has a fixed number of experiments in mind, then DP provides improved results because it takes into account experiments over the full horizon, but for the same reason, it can be disadvantageous if one is interested in stopping experimentation once MOCU reduction falls below a given threshold, meaning that further experimentation is not worth the cost. Although our focus in this article has been intervention in GRNs, it should be recognized that Greedy-MOCU has been used for optimal experimental design in other environments such as the development of optimally performing materials[34] and optimal signal filtering.[35] Dynamic programming can be applied for these problems in the same way it has been done here.

Author Contributions

MI and RD contributed in developing method and also the initial manuscript. ERD and UMB-N contributed ideas for the methods and also revision of the manuscript. All authors submitted comments, read and approved the final manuscript.

REFERENCES

1. Shmulevich I, Dougherty ER, Kim S, Zhang W. Probabilistic Boolean networks: a rule-based uncertainty model for gene regulatory networks. *Bioinformatics*. 2002;18:261–274.
2. Imani M, Braga-Neto U. Control of gene regulatory networks with noisy measurements and uncertain inputs. *IEEE T Contr Netw Syst*. 2018;5:760–769.
3. Imani M, Braga-Neto U. Finite-horizon LQR controller for partially-observed Boolean dynamical systems. *Automatica*. 2018;95:172–179.
4. Shmulevich I, Dougherty ER, Zhang W. Control of stationary behavior in probabilistic Boolean networks by means of structural intervention. *J Biol Syst*. 2002;10:431–445.
5. Dougherty ER, Pal R, Qian X, Bittner ML, Datta A. Stationary and structural control in gene regulatory networks: basic concepts. *Int J Syst Sci*. 2010;41:5–16.
6. Xiao Y, Dougherty ER. The impact of function perturbations in Boolean networks. *Bioinformatics*. 2007;23:1265–1273.
7. Shmulevich I, Dougherty ER, Zhang W. From Boolean to probabilistic Boolean networks as models of genetic regulatory networks. *P IEEE*. 2002;90:1778–1792.
8. Qian X, Dougherty ER. Effect of function perturbation on the steady-state distribution of genetic regulatory networks: optimal structural intervention. *IEEE T Signal Proces*. 2008;56:4966–4976.
9. Bellman R, Kalaba R. Dynamic programming and adaptive processes. *IEEE T Automat Contr*. 1960;5:5–10.
10. Silver TA. Markovian decision processes with uncertain transition probabilities or rewards. 1962. Technical report. DTIC document.
11. Gozzolino JM, Gonzalez-Zubieta R. Markovian decision processes with uncertain transition probabilities. 1965. Technical report. DTIC document.
12. Kuznetsov VP. Stable detection when the signal and spectrum of normal noise are inaccurately known. *Telecomm Radio Eng*. 1976;30:58–64.
13. Kassam SA, Lim TL. Robust Wiener filters. *J Frank Inst*. 1977;304:171–185.
14. Poor H. On robust Wiener filtering. *IEEE T Automat Contr*. 1980;25:531–536.
15. Grigoryan AM, Dougherty ER. Bayesian robust optimal linear filters. *Signal Process*. 2008;81:2503–2521.
16. Dalton LA, Dougherty ER. Intrinsically optimal Bayesian robust filtering. *IEEE T Signal Proces*. 2014;62:657–670.
17. Dalton LA, Dougherty ER. Optimal classifiers with minimum expected error within a Bayesian framework—Part I: discrete and Gaussian models. *Pattern Recogn*. 2013;46:1301–1314.
18. Dalton LA, Dougherty ER. Optimal classifiers with minimum expected error within a Bayesian framework—Part II: properties and performance analysis. *Pattern Recogn*. 2013;46:1288–1300.
19. Yoon BJ, Qian X, Dougherty ER. Quantifying the objective cost of uncertainty in complex dynamical systems. *IEEE T Signal Proces*. 2013;61:2256–2266.
20. Lindley DV. On a measure of the information provided by an experiment. *Ann Math Stat*. 1956;27:986–1005.
21. Raiffa H, Schlaifer RO. *Applied Statistical Decision Theory*. Cambridge, MA: The MIT Press; 1961.
22. Huan X, Marzouk Y. Sequential Bayesian optimal experimental design via approximate dynamic programming. arXiv:1604.08320; 2016.
23. Dehghannasiri R, Yoon BJ, Dougherty ER. Optimal experimental design for gene regulatory networks in the presence of uncertainty. *IEEE/ACM T Comput Biol Bioinform*. 2015;12:938–950.
24. Dehghannasiri R, Yoon BJ, Dougherty ER. Efficient experimental design for uncertainty reduction in gene regulatory networks. *BMC Bioinformatics*. 2015; 16:S2.
25. Mohsenizadeh D, Dehghannasiri R, Dougherty E. Optimal objective-based experimental design for uncertain dynamical gene networks with experimental error. *IEEE/ACM T Comput Biol Bioinform*. 2018;15:218–230.
26. Frazier PI, Powell WB, Dayanik S. A knowledge-gradient policy for sequential information collection. *SIAM J Control Optim*. 2008;47:2410–2439.
27. Kauffman SA. *The Origins of Order: Self-organization and Selection in Evolution*. New York: Oxford University Press; 1993.
28. Imani M, Braga-Neto U. Maximum-likelihood adaptive filter for partially observed Boolean dynamical systems. *IEEE T Signal Proces*. 2017;65: 359–371.
29. Albert R, Othmer HG. The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in *Drosophila melanogaster*. *J Theor Biol*. 2003;223:1–18.
30. Kauffman S, Peterson C, Samuelsson B, Troein C. Random Boolean network models and the yeast transcriptional network. *Proc Natl Acad Sci U S A*. 2003;100: 14796–14799.
31. Fauré A, Naldi A, Chaouiya C, Thieffry D. Dynamical analysis of a generic Boolean model for the control of the mammalian cell cycle. *Bioinformatics*. 2006;22:e124–e131.
32. Bertsekas DP. *Dynamic Programming and Optimal Control*. Vol 1. Belmont, MA: Athena Scientific; 1995.
33. Lau KY, Ganguli S, Tang C. Function constrains network architecture and dynamics: a case study on the yeast cell cycle Boolean network. *Phys Rev E*. 2007; 75:051907.
34. Dehghannasiri R, Qian X, Dougherty ER. Optimal experimental design in the context of canonical expansions. *IET Signal Process*. 2017;11: 942–951.
35. Dehghannasiri R, Xue D, Balachandran PV, et al. Optimal experimental design for materials discovery. *Comput Mater Sci*. 2017;129:311–322.