

---

# Optimal Control of Network Structure Growth

---

**Dominik Thalmeier**  
Radboud University Nijmegen  
Nijmegen, the Netherlands  
d.thalmeier@science.ru.nl

**Vicenç Gómez**  
Universitat Pompeu Fabra  
Barcelona, Spain  
vicen.gomez@upf.edu

**Hilbert J. Kappen**  
Radboud University Nijmegen  
Nijmegen, the Netherlands  
b.kappen@science.ru.nl

## Abstract

We formulate the problem of influencing the growth of a network as a stochastic optimal control problem in which a structural cost function penalizes undesired topologies. We approximate this control problem with a restricted class of control problems that can be solved using probabilistic inference methods. To deal with the increasing problem dimensionality, we introduce an adaptive importance sampling method for approximating the optimal control. We illustrate this methodology in the context of formation of information cascades, considering the task of influencing the structure of a growing conversation thread, as in Internet forums. Using a realistic model of growing trees, we show that our approach can yield conversation threads with better structural properties than the ones observed without control.

## 1 Introduction

Many complex systems can be characterized by the topology of an underlying network. Examples of such systems are human interaction networks, where the links may represent transmitting opinions [2], habits [5], finances [1] or viruses [15]. Being able to control, or just influence in some way, the dynamics of such complex networks may lead to important progress, for example, avoiding financial crises, preventing epidemic outbreaks or maximizing information spread in marketing campaigns. The control of the dynamics on networks is a very challenging problem that has attracted significant interest recently [14, 3, 6]. Existing approaches typically consider network controllability as the controllability of the dynamical system induced by the underlying network structure. While it is agreed that network controllability critically depends on the network structure, the problem of how to control the network structure itself while it is evolving remains open.

The network structure is determined by the dynamics of addition/deletion of nodes/links over time. Here we address the problem of influencing this dynamics in the framework of stochastic optimal control. We propose an approximation based on a special class of stochastic optimal control problems, known as Kullback-Leibler (KL) control or Linearly-Solvable Markov Decision Problems (LMDPs) [10, 20]. For such problems, one can use adaptive importance sampling methods that scale well in when the state space increases, as in our case, and the standard approach through dynamic programming is no longer feasible. The optimal solution for the KL-control problem tends to be sparse, so that only a few *next* states become relevant, effectively reducing the branching factor of the original problem. We use the obtained solution of the KL-control problem to compute an action in the original problem that does not necessarily belong to the KL-control class.

## 2 Optimal Network Growth as a Kullback-Leibler Control Problem

Let  $x_t \in \mathcal{X}$  denote the growing structure of the network at time-step  $t$ , with  $\mathcal{X}$  the set of all possible network structures. We define the natural (uncontrolled) growth process of the network as a Markov

chain with transition probabilities  $p(x'|x)$ . Our controls directly specify the transition probabilities between two subsequent network structures, e.g.  $u(x'|x, t)$ . At each time-step  $t$ , we incur an arbitrary application-dependent cost  $r(x, t)$ , which is assigned when the state is reached. For example, if one wants to favour networks with large average clustering coefficient  $C(x)$ , then  $r(x, t) = -C(x)$ .

Let the probability of an uncontrolled network trajectory (path)  $x_{t+1:T}$  be  $\bar{p} = p(x_{t+1:T}|x, t) = \prod_{t'=t+1}^{T-1} p(x_{t'+1}|x_{t'})$  and, similarly, denote the probability of a controlled path as  $\bar{u}$ . Define the total expected cost of a controlled path as

$$C_{\text{KL}}^\lambda(x, t, u(\cdot)) = r(x, t) + \left\langle \sum_{t'=t+1}^T r(x_{t'}, t') \right\rangle_{\bar{u}} + \lambda D_{\text{KL}}(\bar{u} \parallel \bar{p}) \quad (1)$$

with the KL-divergence  $D_{\text{KL}}(\bar{u} \parallel \bar{p}) = \left\langle \log \frac{\bar{u}}{\bar{p}} \right\rangle_{\bar{u}}$ , which measures the closeness between path distributions. Parameter  $\lambda$  thus regulates the influence of the control on the natural network dynamics.

The control problem consisting in minimizing  $C_{\text{KL}}^\lambda$  w.r.t. the control  $u(x'|x, t)$  belongs to the KL-control class and has a closed form solution [20, 10]. The probability distribution of an optimal path  $u_{\text{KL}}^*(x_{t+1:T}|x, t)$  that minimizes Eq. (1) is

$$u_{\text{KL}}^*(x_{t+1:T}|x, t) = \frac{p(x_{t+1:T}|x, t) \phi(x_{t+1:T})}{\langle \phi(x_{t+1:T}) \rangle_{p(x_{t+1:T}|x, t)}}, \quad \phi(x_{t+1:T}) := \exp\left(-\frac{1}{\lambda} \sum_{t'=t+1}^T r(x_{t'}, t')\right). \quad (2)$$

Plugging this into Eq. (1) and minimizing gives the optimal cost-to-go for state  $x$  and time  $t$

$$J_{\text{KL}}^\lambda(x, t) = r(x, t) - \lambda \log \langle \phi(x_{t+1:T}) \rangle_{p(x_{t+1:T}|x, t)}, \quad (3)$$

which can be numerically approximated using paths sampled from the natural network growth dynamics  $p(x_{t+1:T}|x, t)$ . The optimal control at time  $t$  is the marginal state-transition distribution

$$u_{\text{KL}}^*(x'|x, t) \propto \sum_{x_{t+2:T}} u_{\text{KL}}^*(x_{t+1} = x', x_{t+2:T}|x, t) = p(x'|x) \exp\left(-\frac{J_{\text{KL}}^\lambda(x', t+1)}{\lambda}\right). \quad (4)$$

This resembles a Boltzmann distribution with temperature  $\lambda$  where the optimal cost-to-go takes the role of an energy. For high values of  $\lambda$ ,  $u_{\text{KL}}^*(x'|x, t)$  deviates only a little from the natural network growth, thus the optimal control has a weak influence on the system. In contrast, for low values of  $\lambda$ , the exponential in Eq. (4) becomes very pronounced for the state(s)  $x'$  with the smallest cost-to-go, suppressing the transition probabilities to suboptimal states  $x'$ . Thus the control has a very strong effect on the process. In the limit of  $\lambda$  going to zero, the controlled process becomes deterministic.

In real applications, our control signal may be constrained in different ways and it may not be possible to directly control the transition dynamics between two network structures. Nevertheless, we can use the optimal cost-to-go  $J_{\text{KL}}^\lambda(x', t+1)$  of Eq. (3) as a proxy of the real optimal cost-to-go and select an action greedily according to an estimate of  $J_{\text{KL}}^\lambda(x', t+1)$ . This is the approach taken in this work.

## 2.1 Adaptive Importance Sampling in Growing Networks

A naive way to sample from the optimal growth process consists in sampling paths from the uncontrolled growth process of the network  $p(x'|x)$  and weight them by their corresponding exponentiated state costs. This method is inefficient, specially for low temperatures, when only a few samples with very large weights contribute to the approximation, resulting in very poor estimates. This is a standard problem in Monte Carlo sampling. The Cross-Entropy (CE) method [4, 11] is an adaptive importance sampling algorithm that incrementally updates a baseline sampling policy, which is more sample efficient than the naive sampling method.

We use the CE in the discrete formulation. Our proposal distribution  $u_\omega(x'|x, t)$ , with parameters  $\omega$ , takes the same Markov process form as the optimal control  $u_{\text{KL}}^*$ , Eq. (4). We approximate the cost-to-go by a linear sum of time-dependent feature vectors  $\psi_k(x, t)$  that encode the network structure

$$u_\omega(x'|x, t) \propto p(x'|x) \exp\left(-\frac{\tilde{J}_{\text{KL}}(x', \omega(t))}{\lambda}\right), \quad \tilde{J}_{\text{KL}}(x, \omega(t)) = \sum_k \omega_k(t) \psi_k(x, t). \quad (5)$$

The CE method alternates the following steps until convergence:

**Step 1:** The optimal control  $\bar{u}_{\text{KL}}^* = u_{\text{KL}}^*(x_{t+1:T}|x, t)$  is estimated using sample paths drawn from the parametrized proposal distribution  $u_\omega(x'|x, t)$ . We generate  $M$  sample paths  $x_{t+1:T}^{(i)}, i = 1, \dots, M$  from  $u_\omega(x'|x, t)$  and reweight them with the corresponding importance sampling weights

$$u_{\text{KL}}^*(x_{t+1:T}^{(i)}|x, t) \propto u_\omega(x_{t+1:T}^{(i)}|x, t) \frac{p(x_{t+1:T}^{(i)}|x, t)}{u_\omega(x_{t+1:T}^{(i)}|x, t)} \exp\left(-\frac{1}{\lambda} \sum_{t'=t+1}^T r(x_{t'}^{(i)}, t')\right). \quad (6)$$

**Step 2:** The time-dependent weights  $\omega_k(t)$  of the importance sampler are updated such that  $u_\omega(x'|x, t)$  becomes closer to the optimal sampling distribution. This update involves again a KL-minimization, but with respect to the importance sampling distribution  $\bar{u}_\omega$

$$\operatorname{argmin}_\omega D_{\text{KL}}(\bar{u}_{\text{KL}}^* \parallel \bar{u}_\omega) = \operatorname{argmin}_\omega \left\langle \log \frac{\bar{u}_{\text{KL}}^*}{\bar{u}_\omega} \right\rangle_{\bar{u}_{\text{KL}}^*} = \operatorname{argmin}_\omega - \langle \log \bar{u}_\omega \rangle_{\bar{u}_{\text{KL}}^*} =: -D(\omega), \quad (7)$$

where we dropped the term  $\langle \log \bar{u}_{\text{KL}}^* \rangle_{\bar{u}_{\text{KL}}^*}$ , since  $\omega$  does not depend on it. We minimize Eq. (7) by gradient descent. The parameters  $\omega_k(t)$  are initialized with zeros, which makes the initial proposal distribution equivalent to the uncontrolled dynamics.

### 3 Application: Growing Cascades

We now illustrate the presented framework in the context of growing information cascades. In particular, we focus on the task of influencing the growth of online conversations, that occur, for example, in online forums such as weblogs [12] or news aggregators [7]. Conversation threads start with an initial post and are followed by a cascade of reactions from different users that comment either to the original post or to comments from other users.

We assume an underlying (not observed) community of users and focus on the discussion thread as a growing tree. In this application, we ignore the content of the messages. Since we can not control directly what is the node that will receive the next comment, we propose to use the user interface as a control mechanism to influence indirectly the thread formation process. In our case, the control signal is a recommended comment to which the next user can reply. Our goal is thus to modify the structure of a tree (the conversation) in a certain way while it evolves, by influencing its growth indirectly.

In this application, we are interested in trees with large Hirsch indices (h-index). A tree with h-index  $h$  has  $h$  comments each of which have received at least  $h$  replies, thus measures how distributed the comments of users on previous comments are. We model the problem as a finite horizon task with end-cost only, defined as  $r(x, t) = -\delta_{t,T} \cdot h(x)$ , where  $h(x)$  is the h-index of the tree  $x$ . Since the h-index is a function of the degree sequence of all the nodes in the tree, we use the degree histogram as features  $\psi_k(x, t)$  for the parametrized form of the optimal cost-to-go, Eq. (5). That is, feature  $\psi_k(x, t)$  is the number of nodes with degree  $k$  in the tree  $x$  at time-step  $t$ .

We learn the uncontrolled process  $p(x'|x)$  from a dataset of online discussions from Slashdot<sup>1</sup>. We use a generative model of cascades introduced in [8]. This model determines the probability of a comment to attract a reply by means of an interplay between the *popularity* of a comment (number of replies that a comment has already received), its *novelty* (elapsed time since the comment appeared in the thread) and a *root bias* (certain trendiness of the main post). Such a model has proven to be successful in capturing the structural properties and the temporal evolution of discussion threads present in very diverse platforms [8].

<sup>1</sup>www.slashdot.org

---

**Algorithm 1** Cross-Entropy method for KL-control of network growth

---

**Require:** importance sampler  $u_\omega$ ,  
state  $x$ , feature space  $\psi(\cdot)$ ,  
number of samples  $M$ , learning rate  $\eta$

$l \leftarrow 0$

$\omega_k^{(l)}(t) \leftarrow 0$ , Initialize weights for all  $k, t, l$

$x_{t+1:T}^{(i)} \leftarrow \text{draw } M \text{ paths } \sim u_{\omega^{(l)}}, i = 1, \dots, M$

**repeat**

compute gradients  $\frac{\partial D(\omega^{(l)})}{\partial \omega_k(t)}$

$\omega_k^{(l+1)}(t) \leftarrow \omega_k^{(l)}(t) + \eta \frac{\partial D(\omega^{(l)})}{\partial \omega_k(t)}$  for all  $k, t, l$

$x_{t+1:T}^{(i)} \leftarrow \text{draw } M \text{ samples } \sim \tilde{u}_{\omega^{(l+1)}}$

$l \leftarrow l + 1$

**until** convergence

**return** estimate  $u_{\text{KL}}^*(x'|x, t)$

---

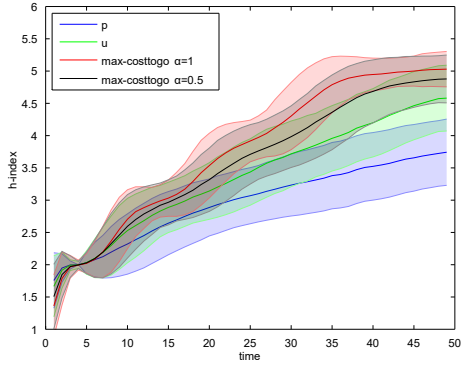


Figure 1: **h-index of a growing cascade:** (blue) the uncontrolled dynamics according to the learned model, (green) the KL-optimally controlled approximation and our action selection based control for  $\alpha = 0.5$  (black) and  $\alpha = 1$  (red). In all cases,  $\lambda = 0.2$  and  $M = 10^3$  paths.

As expected, we observe a faster increase, on average, than using the uncontrolled dynamics. The red and black curves show the evolution of the h-index using our proposed highlighting mechanism, with the expected cost-to-go  $J_{\text{KL}}^\lambda$  of the KL-optimal control with  $\lambda = 0.2$ , for  $\alpha = 1$  and  $\alpha = 0.5$ , respectively. In both cases the obtained h-index is even higher than the one obtained with the KL-control approximation. Therefore, the objective for this task, to increase the h-index, can be achieved through our action selection strategy. As expected, a higher value of  $\alpha = 1$  leads to higher h-indices than a lower one  $\alpha = 0.5$ .

Figure 2 shows examples of discussion threads. The one resulting from applying our action selection strategy has h-index 6 while the data and model have both h-index 4.

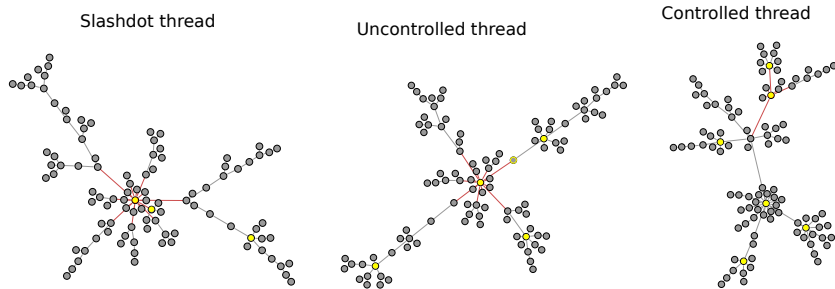


Figure 2: **(left)** A thread from the data (Slashdot), **(middle)** an uncontrolled thread generated from the model and **(right)** a controlled thread. In yellow, the nodes that contribute to the h-index.

## 4 Conclusions

We have presented a method for controlling the structure of a growing network using stochastic optimal control. Our approach is inspired in recent developments on optimal control with information-processing constraints [19, 18, 16, 13, 9]. The KL-control formulation effectively introduces a regularizer which penalizes deviations from the natural network growth process. One advantage of this approach is that the optimal control can be solved by adaptive importance sampling.

We have illustrated the effectiveness of our method on the task of influencing the growth of conversation cascades. This is a non-trivial task characterized by a sparse, delayed reward, since the h-index remains constant during most of the time, and therefore a greedy strategy is not possible.

To evaluate the proposed framework we use a simulated environment, without real users. We set the horizon time  $T = 50$  and start from a thread with a single node. The state-space consists therefore of  $50! \approx 3^{64}$  states. At each time-step, a new node is added to the thread by a (simulated) user. For that, we first choose which node to highlight (optimal action) using Algorithm 1. We then simulate the user, who either selects the highlighted node with some probability  $p' = \alpha/(1 - \alpha)$  or chooses to ignore it with probability  $1 - p'$ . In that case, the parent node is chosen according to the natural growth process  $p(x'|x)$  learned from Slashdot data. This process is repeated until the end time  $T$ .

Figure 1 shows the evolution of the h-index using different control mechanisms derived from our proposed framework. To our knowledge, there are no current alternative methods to compare with in this complex task. The blue curve corresponds to the uncontrolled dynamics  $p(x'|x)$ . In green, we show the evolution of the h-index using our estimate of  $u_{\text{KL}}^*$ , for temper-

## References

- [1] Hamed Amini, Rama Cont, and Andreea Minca. Resilience to contagion in financial networks. *Mathematical finance*, 26(2):329–365, 2016.
- [2] D. Centola and A. Baronchelli. The spontaneous emergence of conventions: An experimental study of cultural evolution. *Proceedings of the National Academy of Sciences*, 112(7):1989–1994, 2015.
- [3] Sean P. Cornelius, William L. Kath, and Adilson E. Motter. Realistic control of network dynamics. *Nature Communications*, 4(1942), Jun 2013.
- [4] P.-T. De Boer, D. P. Kroese, S. Mannor, and R. Y. Rubinstein. A tutorial on the cross-entropy method. *Annals of operations research*, 134(1):19–67, 2005.
- [5] M. Farajtabar, N. Du, M. Gomez-Rodriguez, I. Valera, H. Zha, and L. Song. Shaping social activity by incentivizing users. In *Advances in neural information processing systems*, pages 2474–2482, 2014.
- [6] J. Gao, Y.-Y. Liu, R. M. D’Souza, and A.-L. Barabási. Target control of complex networks. *Nature communications*, 5(5415), 2014.
- [7] V. Gómez, A. Kaltenbrunner, and V. López. Statistical analysis of the social network and discussion threads in Slashdot. In *Proceedings of the 17th international conference on World Wide Web*, pages 645–654. ACM, 2008.
- [8] V. Gómez, H. J. Kappen, N. Litvak, and A. Kaltenbrunner. A likelihood-based framework for the analysis of discussion threads. *World Wide Web*, 16(5-6):645–675, 2013.
- [9] V. Gómez, H. J. Kappen, J. Peters, and G. Neumann. Policy search for Path-Integral control. In *Machine Learning and Knowledge Discovery in Databases*, pages 482–497. Springer, 2014.
- [10] H. J. Kappen, V. Gómez, and M. Opper. Optimal control as a graphical model inference problem. *Machine Learning*, 87(2):159–182, 2012.
- [11] H. J. Kappen and H. C. Ruiz. Adaptive importance sampling for control and inference. *Journal of Statistical Physics*, 162(5):1244–1266, 2016.
- [12] J. Leskovec, M. McGlohon, C. Faloutsos, N. S. Glance, and M. Hurst. Patterns of cascading behavior in large blog graphs. In *SIAM International Conference on Data Mining*, volume 7, pages 551–556, 2007.
- [13] S. Levine and V. Koltun. Guided policy search. *Proceedings of The 30th International Conference on Machine Learning*, 3:1–9, 2013.
- [14] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási. Controllability of complex networks. *Nature*, 473(7346):167–173, 2011.
- [15] R. Pastor-Satorras and A. Vespignani. Epidemic spreading in scale-free networks. *Physical Review Letters*, 86(14):3200–3203, Apr 2001.
- [16] K. Rawlik, M. Toussaint, and S. Vijayakumar. On stochastic optimal control and reinforcement learning by approximate inference. In *Int. Conf. on Robotics Science and Systems*, 2012.
- [17] D. Thalmeier, V. Gómez, and H. J. Kappen. Action selection in growing state spaces: control of network structure growth. *J. Phys. A: Math. Theor.*, 49, 2016.
- [18] E. Theodorou and E. Todorov. Relative entropy and free energy dualities: Connections to path integral and KL control. In *Decision and Control, 2012 IEEE 51st Annual Conference on*, pages 1466–1473. IEEE, 2012.
- [19] N. Tishby and D. Polani. Information theory of decisions and actions. In *Perception-action cycle*, pages 601–636. Springer, 2011.
- [20] E. Todorov. Efficient computation of optimal actions. *Proceedings of the National Academy of Sciences*, 106(28):11478–11483, 2009.