

Optimizing Estimated Directed Information over Discrete Alphabets

Dor Tsur
Ben-Gurion University
dortz@post.bgu.ac.il

Ziv Aharoni
Ben-Gurion University
zivah@post.bgu.ac.il

Ziv Goldfeld
Cornell University
goldfeld@cornell.edu

Haim H. Permuter
Ben-Gurion University
haimp@bgu.ac.il

Abstract—Directed information (DI) is a fundamental measure for the study and analysis of sequential stochastic models. In particular, when optimized over the input distribution, it characterizes the capacity of general communication channels. However, existing optimization methods for discrete input alphabets assume full knowledge of the channel model, and are therefore not applicable when only samples are available. We derive a new method that overcomes this limitation and enables optimizing DI over unknown channels. To that end, we formulate the problem as a Markov decision process and leverage reinforcement learning techniques to optimize a deep generative model of the channel input probability mass function (PMF). Combining our optimizer with the DI neural estimator, we obtain an end-to-end estimation-optimization scheme which is applied for estimating the capacity of various discrete channels with memory. We provide empirical results that demonstrate the utility of the proposed framework and further show how to use the optimized PMF generator to obtain theoretical bounds on the feedback capacity for unifilar finite state channels.

I. INTRODUCTION

Originally proposed for the study of channels with feedback, directed information (DI) quantifies statistical and temporal dependencies between stochastic processes [1]. It has seen a variety of applications in communications [2], [3], portfolio theory [4], computational biology [5], neuroscience [6] and machine learning [7], [8]. Oftentimes, one wishes to optimize the DI with respect to (w.r.t.) the distribution of some of the involved stochastic processes, e.g., channel capacity is given by the maximized DI over input distributions [9].

There are several existing DI optimization methods. The authors in [10] and [11] propose a generalization of the Blahut-Arimoto algorithm [12], [13] for DI. Another approach formulates the DI optimization as a dynamic program (DP) [14], which is solved via recursive algorithms. The authors of [3], [15] applied this method for binary unifilar finite-state channels (FSCs) with feedback, which are FSCs whose state evolves as a function of the past input, output, and state tuple. This method was generalized to large alphabets via reinforcement learning (RL) [16]. However, when feedback is not present, or the FSC is not unifilar, the resulting objective may not be computable [17]. In such cases, the capacity can be bounded utilizing an auxiliary random process termed Q -graph [18], [19]. These methods apply to a larger class of channels, but require an exhaustive search over all possible Q -graphs. In addition to the mentioned challenges, all methods require full knowledge of the channel probabilistic model.

In this work we propose a new method for DI optimization over discrete input alphabets for any stationary and ergodic channel, treating it as a black box whose outputs can be sampled. To account for the lack of channel knowledge, we employ the DI neural estimator (DINE) [20]. The DINE is based on the Donsker-Varadhan representation, where optimal potentials are parametrized by recurrent neural networks (RNNs), expectations are approximated by samples means, and the resulting parametric-empirical objective is optimized over the parameter space. In [20], the DINE was used to compute the capacity of continuous-input channels by coupling it with an optimizer over continuous distributions. This approach utilized an auxiliary RNN-based generative model for the input, and jointly optimized it with DINE by propagating gradients through both models. For memoryless channels, additional joint estimation-optimization methods were proposed over continuous inputs spaces [21], [22]. However, when the input space is discrete, end-to-end differentiability fails to hold. This calls for new ideas to treat channels with discrete alphabets.

We propose herein a new method for optimizing DINE over discrete input alphabets. We model the input distribution by an RNN-based probability mass function (PMF) generator. An optimization objective for the PMF generator is derived using reinforcement learning (RL) techniques by formulating the problem as a Markov decision process (MDP). Applying the policy gradients theorem [23], along with function approximation results and Monte-Carlo averaging, we obtain a tractable optimization objective that is a function of the PMF generator and the DINE outputs. This yields an optimization-estimation algorithm for computing the capacity of channels with discrete inputs, while assuming access only to channel output samples. We apply the proposed method to several channels with memory, and show that for all considered channels the method either achieves the theoretical capacity value or converges between known upper and lower bounds. We then demonstrate how the optimized PMF generator structure coincides with previously known capacity-achieving encoding schemes. We also employ the generator to estimate a Q -graph [24], which can be plugged into the algorithm from [18] to obtain tight bounds on the feedback capacity of unifilar FSCs.

II. PROBLEM SETUP AND MDP FORMULATION

We present the DI optimization method in the context of channel capacity estimation. Channel capacity, both with and

without feedback, is characterized via DI as [1], [9]

$$C = \lim_{n \rightarrow \infty} \sup_P \frac{1}{n} I(X^n \rightarrow Y^n), \quad (1)$$

where $I(X^n \rightarrow Y^n) := \sum_{i=1}^n I(X^i; Y_i | Y^{i-1})$ is the DI. For feedforward channels (i.e., without feedback), P is of the form P_{X^n} . When feedback is present, P is taken as a causal conditioned distribution $P_{X^n | Y^{n-1}} := \prod_{i=1}^n P_{X_i | X^{i-1} Y^{i-1}}$. We assume that the channel model is stationary and unknown, and $|\mathcal{X}| = k < \infty$. For simplicity of presentation, we focus on discrete input and output channels and describe them by a conditional PMF $p_{Y^n | X^n}$. Nevertheless, the proposed method readily extends to arbitrary channels with discrete inputs.

A. PMF Generator and Optimization Objective

To generate the input PMF and optimize the DI rate w.r.t. it, we propose a deep generative model that takes an input-output pair from $\mathcal{X} \times \mathcal{Y}$ and a simplex vector (that models the current input PMF), and outputs a new simplex vector (the updated input PMF). The PMF generator corresponding to a parameter vector $\phi \in \mathbb{R}^d$ is denoted by $h_\phi : \mathcal{X} \times \mathcal{Y} \times \Delta_k \mapsto \Delta_k$, where Δ_k is the k -dimensional probability simplex. The model is randomly initialized at some p_0^ϕ , from which we sample X_0^ϕ and pass it through the channel to obtain Y_0 .¹ The t th PMF is defined iteratively, as

$$p_t^\phi := h_\phi(X_{t-1}^\phi, Y_{t-1}, p_{t-1}^\phi), \quad t \geq 1,$$

where $(X_t^\phi, Y_t) \sim p_t^\phi p_{Y_t | X_t Y^{t-1}}$. Note that for fixed ϕ , the input-output history $(X^{\phi, t-1}, Y^{t-1})$ determines the value of p_t^ϕ . To emphasize this dependence, with some abuse of notation, we also write $p_t^\phi(\cdot) = p_t^\phi(\cdot | X^{\phi, t-1}, Y^{t-1})$. Our goal is to optimize h_ϕ over ϕ to maximize the corresponding DI. To that end, we seek a tractable expression for the DI rate gradient $\nabla_\phi (\lim_{n \rightarrow \infty} \frac{1}{n} I(X^{\phi, n} \rightarrow Y^n))$. We formulate the DI rate optimization as an MDP and employ RL techniques to arrive at an objective whose gradients coincide with the above.

B. MDP Formulation

MDPs model sequential decision problems with recursive relations between the involved variables [14]. We consider an infinite-horizon average-reward MDP, i.e., where the objective is given by

$$\rho(\pi) := \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \mathbb{E} [r(U_t, Z_t)], \quad (2)$$

with Z_t being the MDP state, U_t the action drawn according to policy π , and r the reward function. For simplicity, we exploit the stationarity of the model and apply a reverse time shift operator on each timestep, such that the most recent step remains $t = 0$ throughout.

Mapping the DI rate optimization to an MDP allows utilizing RL methods and arriving at a tractable optimization objective for h_ϕ . To do so, view the channel input generator

¹We note that Y_0 also depends on ϕ through X_0^ϕ , but chose to suppress this dependence in our notation for simplicity.

TABLE I
DI OPTIMIZATION MDP FORMULATION

MDP	DI optimization
State Z_t	X_{-t}^{-1}, Y_{-t}^{-1}
Action U_t	X_0
Disturbance W_t	Y_0
Reward $r(U, Z)$	Eq. (3)

as an agent whose action X_0^ϕ at each step is drawn from the parametric policy $\pi_\phi(U_t, Z_t) = p_t^\phi$. The state is taken as the accumulated past channel inputs and outputs, while the disturbance is the channel output, distributed according to $p_{Y_0 | Y_{-t}^{-1}, X_{-t}^0}$. The immediate reward is given by the conditional expectation

$$r(U, Z) = \mathbb{E} \left[\log \frac{p_{Y_0 | Y_{-t}^{-1}, X_{-t}^0}(Y_0 | Y_{-t}^{-1}, X_{-t}^0)}{p_{Y_0 | Y_{-t}^{-1}}(Y_0 | Y_{-t}^{-1})} \middle| X_{-t}^{\phi, 0}, Y_{-t}^{-1} \right]. \quad (3)$$

The formulation is given in Table I. In the following we state the desired equivalence

Theorem 1 (MDP formulation) *The DI rate optimization problem defines an infinite-horizon average-reward MDP, such that $\rho(\pi_\phi) = \lim_{n \rightarrow \infty} \frac{1}{n} I(X^{\phi, n} \rightarrow Y^n)$.*

The proof shows that $Z_t = f(Z_{t-1}, U_t, W_t)$ for some function f , and that $(Z_{-\infty}^{t-1}, U_{-\infty}^{t-1}, W_{-\infty}^{t-1}) - (Z_t, U_t) - W_t$ forms a Markov chain. A key virtue of the MDP formulation is that the optimal policy is guaranteed to be time-invariant [14], motivating the use of RNNs for its parametric approximation.

III. OPTIMIZER DERIVATION

With the MDP formulation above, we next derive a simplified expression for $\nabla_\phi \rho(\pi_\phi)$. We begin by applying the policy gradients theorem [23, Theorem 1]

Theorem 2 (Policy gradients) *Let π_ϕ be a parametric model with parameters ϕ ; let $\rho(\pi_\phi)$ defined as in (2) and d^{π_ϕ} be the states' stationary distribution; define the Q -function*

$$Q^{\pi_\phi}(u, z) := \sum_{t=1}^{\infty} \mathbb{E} [r(U_t, Z_t) - \rho(\pi_\phi) | Z_0 = z, U_0 = u]. \quad (4)$$

Then, for any MDP:

$$\nabla_\phi \rho(\pi_\phi) = \sum_z d^{\pi_\phi}(z) \sum_u \nabla_\phi \pi_\phi(u, z) Q^{\pi_\phi}(u, z). \quad (5)$$

Applying the identity $\partial_x \log f(x) = \frac{\partial_x f(x)}{f(x)}$ to (5) we obtain:

$$\begin{aligned} \nabla_\phi \rho(\pi_\phi) &= \sum_z d^{\pi_\phi}(z) \sum_u \pi_\phi(u, z) \nabla_\phi \log \pi_\phi(u, z) Q^{\pi_\phi}(u, z) \\ &= \mathbb{E} [\nabla_\phi \log \pi_\phi(U, Z) Q^{\pi_\phi}(U, Z)]. \end{aligned} \quad (6)$$

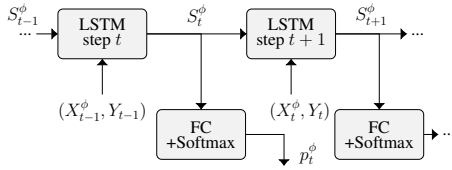


Fig. 1. The PMF model unrolled for feedback capacity.

To compute Q^{π_ϕ} exactly, one must be able to calculate (3), and therefore, know the channel model. We circumvent this by using the DINE [20]. We now present the DINE and then demonstrate its utility for the estimation of Q^{π_ϕ} .

A. Directed Information Neural Estimator

In the proposed formulation the state space is infinite, and therefore the optimization cannot be solved by simple tabular RL schemes. To this end, we incorporate function approximation into (6) using the DINE. Given a sample $D_n = (X^n, Y^n) \sim p_{X^n Y^n}$, the DINE objective is given by

$$\hat{l}(D_n, \theta_y, \theta_{xy}) = \hat{D}_{Y\|X}(D_n, \theta_{xy}) - \hat{D}_Y(D_n, \theta_y), \quad (7)$$

where

$$\begin{aligned} \hat{D}_Y(D_n, \theta_y) &:= \frac{1}{n} \sum_{i=1}^n g_{\theta_y}(Y_i | Y^{i-1}) \\ &\quad - \log \left(\frac{1}{n} \sum_{i=1}^n e^{g_{\theta_y}(\tilde{Y}_i | Y^{i-1})} \right), \\ \hat{D}_{Y\|X}(D_n, \theta_{xy}) &:= \frac{1}{n} \sum_{i=1}^n g_{\theta_{xy}}(Y_i | Y^{i-1}, X^i) \\ &\quad - \log \left(\frac{1}{n} \sum_{i=1}^n e^{g_{\theta_{xy}}(\tilde{Y}_i | Y^{i-1}, X^i)} \right), \end{aligned}$$

g_{θ_y} and $g_{\theta_{xy}}$ are RNNs² with parameters $\theta_y \in \Theta_y \subset \mathbb{R}^{d_y}$ and $\theta_{xy} \in \Theta_{xy} \subset \mathbb{R}^{d_{xy}}$, and $\tilde{Y} \sim \text{Unif}(\mathcal{Y})$. Here, Θ_y and Θ_{xy} are compact subsets of the corresponding Euclidean spaces. The DINE is given by optimization of (7) as

$$\hat{l}(D_n) = \sup_{\theta_{xy} \in \Theta_{xy}} \inf_{\theta_y \in \Theta_y} \hat{l}(D_n, \theta_y, \theta_{xy}). \quad (8)$$

As argued in [20, Theorem 2], the DINE results in a consistent estimator of the DI rate.

B. Q-Function Estimation via DINE

Fix the PMF generator parameters $\phi \in \Phi \subset \mathbb{R}^d$ for a compact parameter space Φ , and consider a dataset $D_n^\phi = (X^{\phi, n}, Y^n) \sim \prod_{t=1}^n p_t^\phi p_{Y_t | Y^{t-1} X^t}$ drawn from this input PMF and the channel. We use the DINE based on D_n^ϕ to approximate Q^{π_ϕ} . Assuming $(g_{\theta_y}, g_{\theta_{xy}})$ are sufficiently close to the supremum-achieving RNNs of (7), we take $\hat{r}_\theta := g_{\theta_{xy}} - g_{\theta_y}$ as a proxy of r , where $\theta = (\theta_y, \theta_{xy})$. To obtain a computable optimization objective we replace the expectation (6) with

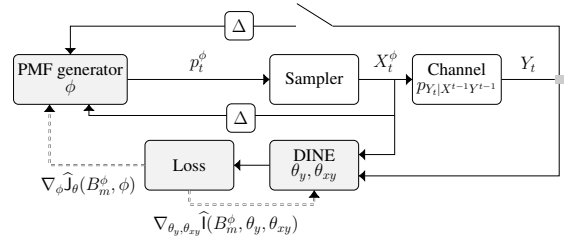


Fig. 2. The complete estimation-optimization model. Dashed arrows represent gradient propagation and filled blocks represent computational models.

a sample mean and take a finite time-horizon T in (4). Accordingly, for fixed θ , we define

$$\hat{J}_\theta(D_n^\phi, \phi) := \frac{1}{n-T} \sum_{t=1}^{n-T} \log p_t^\phi(X_t^\phi) \hat{Q}_{\theta, t}(\theta, D_n^\phi), \quad (9)$$

where the approximated Q -function is given by

$$\hat{Q}_{\theta, t}(\theta, D_n^\phi) := \sum_{i=t}^{t+T-1} \hat{r}_\theta(Y^i, X^{\phi, i}) - \hat{l}(D_n^\phi, \theta_y, \theta_{xy}). \quad (10)$$

When feedforward capacity is considered, the MDP formulation remains unchanged, but we restrict the optimization over policies which are not a function of past channel outputs.

IV. IMPLEMENTATION AND COMPLETE SYSTEM

Recall that channel input PMF p_t^ϕ is obtained as the output of the parametric model h_ϕ , i.e., $p_t^\phi = h_\phi(X_{t-1}^\phi, Y_{t-1}, p_{t-1}^\phi)$. We implement h_ϕ using a long short-term memory (LSTM) network. The LSTM can be stacked with additional fully-connected (FC) layers to increase its expressiveness. The output layer is an $|\mathcal{X}|$ -dimensional softmax layer, denoted by σ_{sm} , which generates outputs that reside in Δ_k . Therefore, the model output (namely, the PMF p_t^ϕ) evolution is described as

$$S_t^\phi = G_1^\phi(X_{t-1}^\phi, Y_{t-1}, S_{t-1}^\phi), \quad p_t^\phi = \sigma_{\text{sm}}(G_2^\phi(S_t^\phi)), \quad (11)$$

where G_1^ϕ is the LSTM map, S_t^ϕ is its output, G_2^ϕ denotes the FC layers and h_ϕ comprises (G_1^ϕ, G_2^ϕ) . In our experiments, we use an LSTM layer with 50 neurons and 2 FC layers with 100 and 50 neurons each. For feedforward capacity estimation, Y_{t-1} is omitted from (11). We illustrate the implementation of h_ϕ in Fig. 1.

The joint estimation-optimization procedure follows an alternating optimization scheme between $\hat{l}(D_n^\phi, \theta_y, \theta_{xy})$ and $\hat{J}_\theta(D_n^\phi, \phi)$. At each iteration, we select a single model for parameters update, while fixing the parameters of the other. We create a batch $B_m^\phi = (X^{\phi, m}, Y^m)$ and a corresponding sequence of PMF generator outputs $p^{\phi, m}$. We pass B_m^ϕ through the DINE model to obtain the corresponding RNN outputs. We then calculate the loss function of the model to be updated and optimize the parameters using stochastic gradient-ascent. The complete system is illustrated in Fig. 2. After a convergence criterion is met or a fixed number of iterations are performed, we evaluate (7) on a long sequence to obtain a numerical

²See [25] for the definition of the RNN class.

Algorithm 1 Discrete alphabet DI optimization and estimation**input:** Discrete channel, feedback indicator**output:** $\hat{l}(D_n^\phi, h_\phi)$, optimized h_ϕ .Initialize $g_{\theta_y}, g_{\theta_{xy}}, h_\phi$ with parameters θ_y, θ_{xy} and ϕ .**if** feedback indicator **then**Add feedback link to h_ϕ **repeat**Compute $(B_m^\phi, p^{\phi, m})$ **if** training DINE **then**Compute $\hat{D}_Y(B_m^\phi, \theta_y), \hat{D}_{Y\|X}(B_m^\phi, \theta_{xy})$

Update DINE parameters:

 $\theta_y \leftarrow \theta_y + \nabla_{\theta_y} \hat{D}_Y(B_m^\phi, \theta_y)$ $\theta_{xy} \leftarrow \theta_{xy} + \nabla_{\theta_{xy}} \hat{D}_{Y\|X}(B_m^\phi, \theta_{xy})$ **else** (Train PMF generator)Compute $\hat{J}_\theta(B_m^\phi, \phi)$

Update PMF generator parameters:

 $\phi \leftarrow \phi + \nabla_\phi \hat{J}_\theta(B_m^\phi, \phi)$ **until** convergenceMonte Carlo evaluation of $\hat{l}(D_n^\phi)$ **return** $\hat{l}(D_n^\phi)$ and h_ϕ

estimate of the channel capacity. The procedure is summarized in Algorithm 1 and its implementation is available on [GitHub](#).

V. PERFORMANCE ANALYSIS

We analyze the performance of the capacity estimator from Algorithm 1 by comparing estimates with known theoretical solutions and/or bounds. We also investigate the structure of the optimized PMF generator and show how to use it for calculating upper and lower bounds on the feedback capacity. We stress that all reference methods presented in this section assume full knowledge of the channel's probabilistic structure.

A. Channel Capacity Estimation

We examine both feedforward and feedback scenarios. For the feedforward capacity we first consider the Gilbert-Elliott (GE) channel [26]—a time-varying BSC whose channel parameter is determined by a latent state that evolves according to a stationary Markov chain. The chain transits between states 'Good' and 'Bad', with transition probabilities b and g from 'Good' to 'Bad' and vice versa. While it is straightforward to show that the optimal input is an i.i.d. $\text{Ber}(0.5)$ process, the GE capacity is determined by a limit expression [27]. We apply the proposed scheme to estimate the capacity of the GE channel and compare our results with a consistent estimation method that relies on prior knowledge of the optimal input distribution [28]. As seen in Fig. 3, our method achieves the capacity for a variety of (b, g) values.

Second, we consider the Trapdoor channel, which is a unifilar FSC, whose state evolves according to $S_t = S_{t-1} \oplus Y_t \oplus X_t$, where $\oplus$ is the XOR operation and its output is equal to S_{t-1} or X_t with equal probability. The feedforward capacity of the Trapdoor channel is an open problem. We apply the proposed algorithm for the feedforward Trapdoor channel and compare

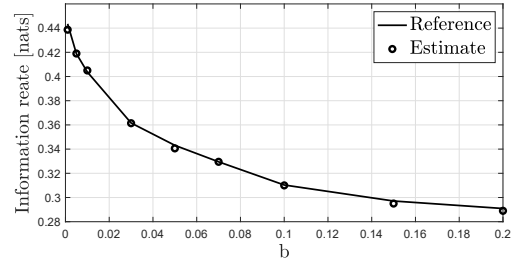


Fig. 3. Estimated capacity of the GE channel versus b (with $g = 3b$), compared to estimates obtained via the method from [28].

with currently known bounds. The results are shown in Fig. 4, where the upper bound is obtained by minimizing the dual capacity upper bound from [19], while the lower bound is obtained numerically from a Blahut-Arimoto-type algorithm [29]. Evidently, our method quickly converges between those bounds, providing a new estimate for the capacity of the Trapdoor channel. By averaging over several simulations, we arrive at an estimate of $\hat{C}_{\text{Trapdoor}} \approx 0.57246$.

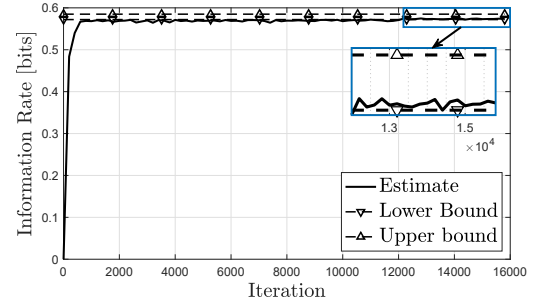


Fig. 4. DINE loss convergence on the feedforward Trapdoor channel, compared with upper and lower bounds from [19] and [29], respectively.

For feedback capacity we consider the Noisy Output is the SState (NOST) channel [30]. The NOST channel is a binary FSC whose state is a stochastic function of the previous channel output. We consider the special case from [30], in which S_t is given as the output of a Z-channel with parameter η and the channel model is then determined from the value of S_t . We estimate the NOST channel capacity for several values of η and compare with the convex optimization approach presented in [30]. As seen in Fig. 5, our method achieves the channel capacity for various values of η .

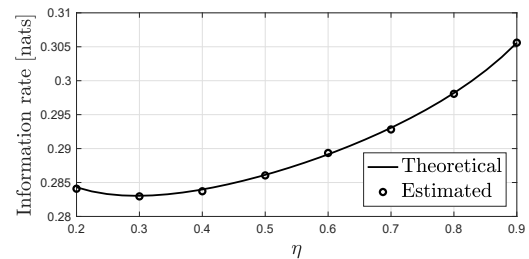


Fig. 5. Estimated capacity for the NOST channel versus flip probability η , compared with the optimization method in [30].

B. Optimized PMF Generator Structure

Consider the binary Ising channel [31], an FSC whose state evolves according to $S_{t+1} = X_t$ and whose outputs are given by either a symmetric Z- or S-channel, depending on S_t . The feedback capacity of this channel was calculated in [15] via dynamic programming techniques. We estimate the feedback capacity using our method and evaluate it in several ways.

1) *Capacity Estimate*: Fig. 6 shows that our method converges to the ground truth capacity after a relatively small number of iterations.

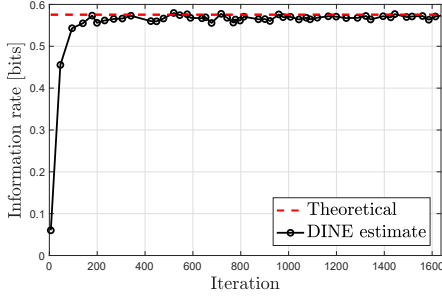


Fig. 6. DINE loss convergence on the Ising channel with feedback.

2) *Learned PMF*: When the joint training process is finished we construct a long trajectory $\{p_t^\phi, x_t, s_t, y_t\}_{i=1}^n$, perform k -means clustering, and analyse the evolution of clustered p_t^ϕ according to past $(x^{t-1}, y^{t-1}, s^{t-1})$. In this case, $p_t^\phi \in [0, 1]$ and is treated as the parameter of a Bernoulli distribution. The transition of p_t^ϕ is presented in Fig. 7. In [15], a feedback capacity-achieving encoding scheme is proposed for the Ising channel. The scheme transmits a binary sequence with a bit flip probability of $\alpha = 0.4503$, with repeated transmissions depending on the feedback value. The learned scheme has the same transition structure as that from [15], but with a slightly different flip probability of 0.456.

3) *Feedback Capacity Bounds*: We employ the optimized PMF generator to calculate bounds on the feedback capacity of the Ising channel. Due to space limitations, we only outline the method, with the full details reserved to an extended version of this work. We leverage the optimization procedure proposed in [18], which is based on Q -graphs. Q -graphs, denoted Q_g , are a function of the channel outputs and are described by a graph with some $|Q_g| < \infty$ number of nodes and $|\mathcal{Y}|$ outgoing edges from each node. Each Q -graph gives rise to upper and lower bounds on the feedback capacity of unifilar FSCs by solving a corresponding minimization or maximization problem over a certain class of input distributions (cf. [18]). As h_ϕ induces a distribution on the channel outputs, we can use it to estimate Q -graphs and explore corresponding capacity bounds.

We focus on estimating $Q_g = p_{S_t|Y^t}$, as its structure is known for the optimal input distribution of the Ising channel [15]. The estimation is performed via a supervised regression over a one-hot representation of the channel states with the cross entropy loss function. We perform k -means and present the resulting Q -graph in Fig. 8. The estimated structure of $p_{S_t|Y^t}$ coincides with the one obtained from the

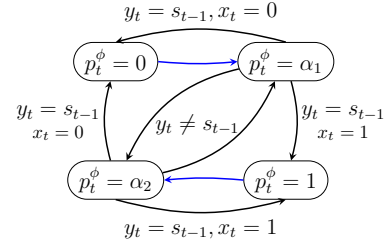


Fig. 7. The learned PMF model, $(\alpha_1, \alpha_2) = (0.456, 0.570)$. A blue arrow denotes a transition that occurs for any value of (x_t, y_t, s_t) .

analytical solution in [15, Theorem 2], although the values of the nodes of our Q -graph slightly differ from those in [15]. Nevertheless, these values do not affect the bound computation [18], and using Equations (13) and (16) from [18] we obtain the following capacity upper and lower bounds:

$$\hat{C}_{\text{Lower}} = 0.5755213, \quad \hat{C}_{\text{Upper}} = 0.5755215.$$

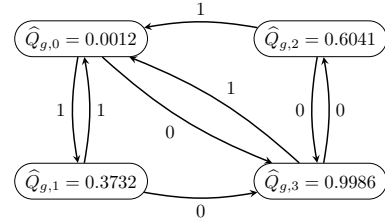


Fig. 8. Learned Q -graph for the Ising channel. Labels denote the value of Y

VI. CONCLUDING REMARKS AND FUTURE WORK

This work proposed a joint estimation-optimization algorithm of DI over discrete input alphabets, leveraging RL techniques and the DINE [20]. The algorithm entails an alternating optimization between the DINE and a deep generative model for the channel input PMF. We employed the algorithm to estimate the feedforward and feedback capacities of a variety of channels with memory and analyzed its performance. Our results demonstrated the high accuracy of the proposed approach in estimating the capacity and recovering the corresponding optimal input distribution. Lastly, the learned input PMF, along with the Q -graph-based optimization procedure of [18], were leveraged to calculate upper and lower bounds on the feedback capacity of unifilar FSCs.

The developed algorithm is distinctive in two key ways: (i) it treats the channel as a black box, relying only on the ability to sample its output, thus, expanding the collection of channels whose capacity can be numerically computed; (ii) it is scalable in memory and dimension. Going forward, we plan to develop theoretical guarantees for the accuracy of the proposed scheme. We also plan to generalize the proposed method to multi-user channels with memory by exploiting multi-agent RL techniques. Our end goal is to create a general framework for capacity estimation that is scalable, efficiently computable, and applicable to a broad scope of channels.

REFERENCES

- [1] James Massey. Causality, feedback and directed information. In *Proc. Int. Symp. Inf. Theory Applic. (ISITA-90)*, pages 303–305. Citeseer, 1990.
- [2] Gerhard Kramer. *Directed information for channels with feedback*, volume 11. Citeseer, 1998.
- [3] Haim Permuter, Paul Cuff, Benjamin Van Roy, and Tsachy Weissman. Capacity of the trapdoor channel with feedback. *IEEE Transactions on Information Theory*, 54(7):3150–3165, 2008.
- [4] Haim H Permuter, Young-Han Kim, and Tsachy Weissman. On directed information and gambling. In *Proceedings of 2008 IEEE International Symposium on Information Theory*, pages 1403–1407. IEEE, 2008.
- [5] Arvind Rao, Alfred O Hero, David J States, and James Douglas Engel. Inference of biologically relevant gene influence networks using the directed information criterion. In *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, volume 2, pages II–II. IEEE, 2006.
- [6] Michael Wibral, Raul Vicente, and Joseph T Lizier. *Directed information measures in neuroscience*. Springer, 2014.
- [7] Yuxun Zhou and Costas J Spanos. Causal meets submodular: Subset selection with directed information. In *Advances In Neural Information Processing Systems*, pages 2649–2657, 2016.
- [8] Stas Tiomkin and Naftali Tishby. A unified bellman equation for causal information and value in markov decision processes. *arXiv preprint arXiv:1703.01585*, 2017.
- [9] Young-Han Kim. A coding theorem for a class of stationary channels with feedback. *IEEE Transactions on Information Theory*, 54(4):1488–1499, 2008.
- [10] Iddo Naiss and Haim H Permuter. Extension of the blahut–arimoto algorithm for maximizing directed information. *IEEE Transactions on Information Theory*, 59(1):204–222, 2012.
- [11] Charalambos D Charalambous and Photios A Stavrou. Directed information on abstract spaces: Properties and variational equalities. *IEEE Transactions on Information Theory*, 62(11):6019–6052, 2016.
- [12] Richard Blahut. Computation of channel capacity and rate-distortion functions. *IEEE transactions on Information Theory*, 18(4):460–473, 1972.
- [13] Suguru Arimoto. An algorithm for computing the capacity of arbitrary discrete memoryless channels. *IEEE Transactions on Information Theory*, 18(1):14–20, 1972.
- [14] Dimitri P Bertsekas et al. *Dynamic programming and optimal control: Vol. 1*. Athena scientific Belmont, 2000.
- [15] Ohad Elishco and Haim Permuter. Capacity and coding for the ising channel with feedback. *IEEE transactions on information theory*, 60(9):5138–5149, 2014.
- [16] Ziv Aharoni, Oron Sabag, and Haim Henri Permuter. Reinforcement learning evaluation and solution for the feedback capacity of the ising channel with large alphabet. *arXiv preprint arXiv:2008.07983*, 2020.
- [17] Andrea Grigorescu, Holger Boche, Rafael F Schaefer, and H Vincent Poor. Capacity of finite state channels with feedback: Algorithmic and optimization theoretic properties. *arXiv preprint arXiv:2201.11639*, 2022.
- [18] Oron Sabag, Bashar Huleihel, and Haim H Permuter. Graph-based encoders and their performance for finite-state channels with feedback. *IEEE Transactions on Communications*, 68(4):2106–2117, 2020.
- [19] Bashar Huleihel, Oron Sabag, Haim H Permuter, Navin Kashyap, and Shlomo Shamai. Computable upper bounds on the capacity of finite-state channels. *IEEE Transactions on Information Theory*, 2021.
- [20] Dor Tsur, Ziv Aharoni, Ziv Goldfeld, and Haim Permuter. Neural estimation and optimization of directed information over continuous spaces. *arXiv preprint arXiv:2203.14743*, 2022.
- [21] Nunzio A Letizia and Andrea M Tonello. Capacity-driven autoencoders for communications. *IEEE Open Journal of the Communications Society*, 2021.
- [22] Farhad Mirkarimi, Stefano Rini, and Nariman Farsad. Neural capacity estimators: How reliable are they? *arXiv preprint arXiv:2111.07401*, 2021.
- [23] Richard S Sutton, David A McAllester, Satinder P Singh, Yishay Mansour, et al. Policy gradient methods for reinforcement learning with function approximation. In *NIPS*, volume 99, pages 1057–1063. Citeseer, 1999.
- [24] Oron Sabag, Haim H Permuter, and Henry D Pfister. A single-letter upper bound on the feedback capacity of unifilar finite-state channels. *IEEE Transactions on Information Theory*, 63(3):1392–1409, 2016.
- [25] Liang Jin, Madan M Gupta, and Peter N Nikiforuk. Universal approximation using dynamic recurrent neural networks: discrete-time version. In *Proceedings of ICNN’95-International Conference on Neural Networks*, volume 1, pages 403–408. IEEE, 1995.
- [26] Edgar N Gilbert. Capacity of a burst-noise channel. *Bell system technical journal*, 39(5):1253–1265, 1960.
- [27] Mordechai Mushkin and Israel Bar-David. Capacity and coding for the gilbert-elliott channels. *IEEE Transactions on Information Theory*, 35(6):1277–1290, 1989.
- [28] Mohammad Rezaeian. Computation of capacity for gilbert-elliott channels, using a statistical method. In *2005 Australian Communications Theory Workshop*, pages 56–61. IEEE, 2005.
- [29] Kingo Kobayashi. Capacity problem of trapdoor channel. In *General Theory of Information Transfer and Combinatorics*, pages 1084–1087. Springer, 2006.
- [30] Eli Shemuel, Oron Sabag, and Haim Permuter. The feedback capacity of noisy output is the state (most) channels, 2021.
- [31] Toby Berger and Flavio Bonomi. Capacity and zero-error capacity of ising channels. *IEEE transactions on information theory*, 36(1):173–180, 1990.