

Cross-domain Transfer Learning and State Inference for Soft Robots via a Semi-supervised Sequential Variational Bayes Framework

Shageenderan Sapai^{1,*}, Junn Yong Loo^{1,*}, Ze Yang Ding², Chee Pin Tan², Raphaël C.-W. Phan¹,
Vishnu Monn Baskaran¹ and Surya Girinatha Nurzaman²

Abstract—Recently, data-driven models such as deep neural networks have shown to be promising tools for modelling and state inference in soft robots. However, voluminous amounts of data are necessary for deep models to perform effectively, which requires exhaustive and quality data collection, particularly of state labels. Consequently, obtaining labelled state data for soft robotic systems is challenged for various reasons, including difficulty in the sensorization of soft robots and the inconvenience of collecting data in unstructured environments. To address this challenge, in this paper, we propose a semi-supervised sequential variational Bayes (DSVB) framework for transfer learning and state inference in soft robots with missing state labels on certain robot configurations. Considering that soft robots may exhibit distinct dynamics under different robot configurations, a feature space transfer strategy is also incorporated to promote the adaptation of latent features across multiple configurations. Unlike existing transfer learning approaches, our proposed DSVB employs a recurrent neural network to model the nonlinear dynamics and temporal coherence in soft robot data. The proposed framework is validated on multiple setup configurations of a pneumatic-based soft robot finger. Experimental results on four transfer scenarios demonstrate that DSVB performs effective transfer learning and accurate state inference amidst missing state labels.

I. INTRODUCTION

Soft robots, often inspired by the morphology of bio-organisms [1], are robots that can bend and conform to their surrounding structures [2]. While this flexibility affords them advantages for deployment in unstructured environments [3], it complicates the development of analytical models due to the inherent nonlinearities and theoretically infinite degrees of freedom [4]. Conversely, data-driven methods present a more accessible alternative to soft robot modelling as the burden of describing the complex system dynamics is discovered from available data. For instance, [5]–[7] have successfully trained data-driven neural networks to address the complexity in modelling the state of soft robotic systems. Nonetheless, these data-driven approaches rely on the availability of high-quality labelled data, representative of the soft robots [8]. Such models demonstrate significant degradation in performance in the absence of labelled data [9].

Unfortunately, data collection is often an arduous process as soft robots must be exhaustively actuated in order to extensively cover their task space [9], [10]. This process is

further challenged in some potential soft robot applications where collection processes are expensive [11]–[14] due to underlying environmental conditions. In addition, soft robots exhibit distinct dynamics under different robot configurations [6], [8] due to their inherent softness and compliance. Data would have to be collected on many configurations to accurately represent a versatile soft robot. Moreover, as opposed to the measurement inputs, collecting relevant state labels is non-trivial [15] and could give rise to missing state labels. For example, the installation of multiple sensors for capturing labelled state data is often avoided to maintain the soft robot’s flexibility. Furthermore, external sensors such as cameras that track the morphological states of the robot are impractical in the unstructured environments within which these robots operate [16]. These challenges inhibit the deployment of data-driven models for soft robots.

To circumvent the requirement of voluminous data, previous studies [5], [17] employ simple or non-parametric models that are less reliant on data abundance but at the cost of having less expressive models. An alternative approach is to generate simulation data to compensate for the lack of real data by employing empirical simulation models. For example, [18], [19] generate simulated data for neural network training via the Finite Element Method (FEM). However, FEM models are generally sensitive to characterizing factors such as varying material properties, inconsistent manufacturing processes, and the influence of external loads [20]. A variation in these factors could result in substantial differences between the simulation models and the actual soft robots [21]. For this reason, data-driven models trained on simulated data often fail to generalize to highly uncertain real-world scenarios [21]. To overcome this, recent works [22]–[26] incorporate sim-to-real transfer strategies to eliminate the gap between the FEM-simulated data and the physical experiment.

Recently, the deep learning paradigm of transfer learning has shown immense potential in alleviating the problem of limited labeled data. Transfer learning approaches aim to adapt the knowledge from some source domain to build data-driven prediction models of good generalization performance on a related target domain with little or no data supervision [15]. In particular, deep learning architectures based on feature encoding, such as autoencoder (AE), have been successfully applied to perform homogeneous transfer learning [27]–[31]. Subsequently, Zheng et al. [32] proposed a transfer learning framework that constructs a universal feature representation via the probabilistic counterpart of AE, variational

*These authors contributed equally.

¹The authors are with School of Information Technology, Monash University Malaysia (E-mail: shageenderan.sapai/loojunyong/raphael.phan/vishnu.mon@monash.edu).

²The authors are with School of Engineering, Monash University Malaysia (E-mail: ding.zeyang/tan.chee.pin/surya.nurzaman@monash.edu).

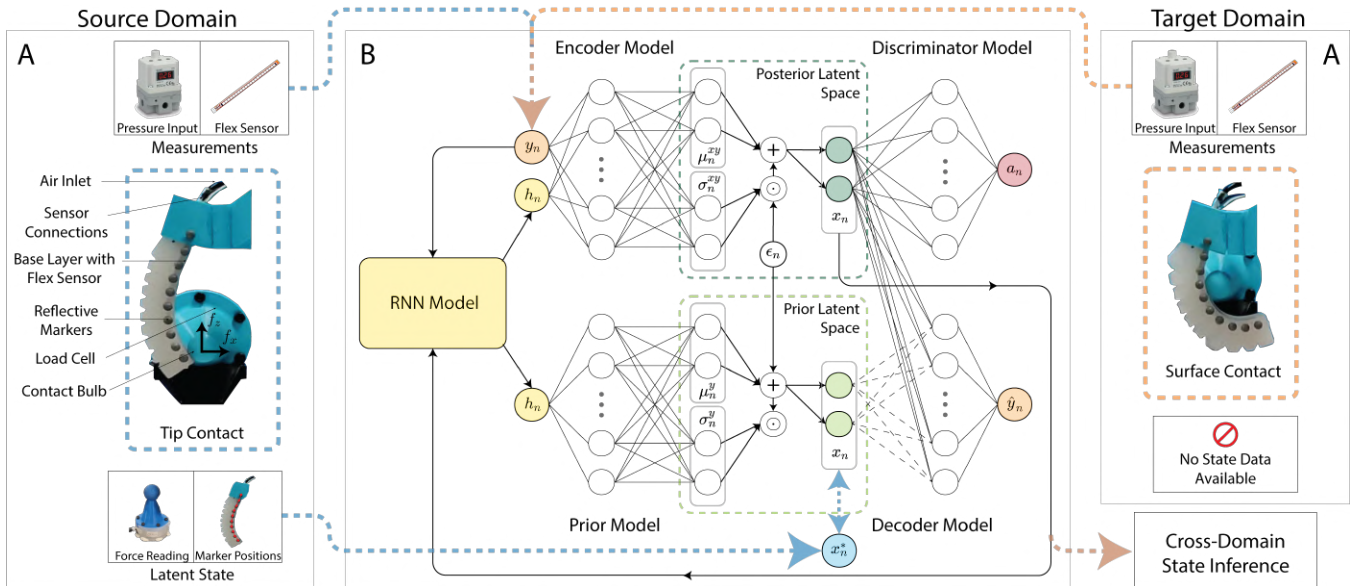


Fig. 1. **A cross-domain transfer learning scenario in soft robotics via the DSVB framework.** (A, Left) Source domain with contact force applied at the PSF tip (i.e., Tip Contact). (A, Right) Target domain with contact forces applied along the PSF surface (i.e., Surface Contact). $f = [f_X, f_Y]^T$ denotes the two-axis contact force. (B) Schematic diagram of the proposed DSVB. The measurements, y_n and posterior latent state, x_n are looped back to the RNN model. The double arrow indicates semi-supervision of the partial state label x_n^* . Coloured dashed lines represent data flow from the source (blue) and target (orange) domains. Dashed (line) flow of the decoder model is used only for online testing. \hat{y}_n denotes the reconstructed measurement.

AE (VAE), which provides a stochastic modelling solution that encapsulates predictive uncertainty. Nonetheless, [32] only considered missing measurement inputs and assumed that the state labels are available. More importantly, these existing AE-based transfer learning approaches rely on the independent and identically distributed (i.i.d.) assumption, which falls short of capturing the rich causal structure and temporal coherence [33] of dynamic soft robot data.

To address the difficulty of acquiring representative soft robot data, in this paper, we propose a domain-adaptable sequential variational Bayes (DSVB) framework for transfer learning and state inference in soft robot systems with missing state labels on certain robot configurations (target domain). Unlike existing AE-based transfer learning approaches that construct internal feature representations without concrete semantics or tangible interpretations, our proposed framework incorporates a semi-supervised disentangle representation learning scheme to endow its latent feature space with the high-level system state representation via partial state labels on the source configurations (source domain). To circumvent the i.i.d. assumption, our proposed DSVB also adopts a recurrent neural network (RNN) to accurately model the nonlinear dynamics of soft robots by constructing a rich representation of the low-level temporal dependencies across the available data. Considering the distinct dynamics and data distributions under different soft robot configurations, a feature space transfer strategy is integrated to promote domain-adaptable latent features that generalize well to both the supervised source and unsupervised target robot domains.

To the best of our knowledge, this is the first work that investigates the feasibility of a probabilistic cross-domain transfer learning framework as an alternative to the costly and laborious soft robot data collection process, alleviating

the immense efforts of acquiring labelled data on many robot configurations. Our contributions are highlighted as follows:

- 1) A DSVB framework is developed to model the complex dynamics of soft robot systems across multiple robot configurations. To facilitate representative state inference, a disentangled representation learning scheme is incorporated to semi-supervise the latent prior distribution on partial state labels from only the source domain, endowing the latent feature space with system state representation of the actual soft robot.
- 2) To achieve cross-domain transfer learning amidst missing state labels on the target domain, a probabilistic feature space transfer strategy is incorporated so that the model knowledge and disentangled representation acquired from the supervised source domain can be adapted to the target domain and guide the unsupervised variational state inference.
- 3) The efficacy of the proposed DSVB is validated on a pneumatic-based soft robot finger (PSF) with multiple configurations. Comparison of results with the state-of-the-art RNNs on four transfer scenarios (e.g., the one shown in Fig. 1) demonstrates that, the DSVB achieves superior state inference performance and effective transfer learning across the supervised source and unsupervised target robot domains.

The remainder of this paper is organized as follows: Section II details the formulation of the proposed DSVB framework. Section III presents a comprehensive case study on the PSF platform. Section IV concludes the paper.

II. METHODS

In this section, a probabilistic DSVB framework is developed to achieve a representative inference model for multi-

configuration soft robots in spite of missing state labels. A schematic diagram of the DSVB is provided in Fig. 1B.

A. Sequential Variational Bayes Objective

In this subsection, we introduce a general sequential Variational Bayes (VB) objective that accounts for the temporal coherence of empirical soft robot data. Variational Bayes (VB) [34] is a class of probabilistic deep learning methods that aims to maximize a lower bound of the joint measurement log-likelihood $\log p_\theta(Y)$. By introducing a latent state variable X , the variational evidence lower bound (ELBO) $\log p_\theta(Y) \leq \mathcal{L}^{\text{ELBO}}(\theta, \vartheta)$ can be obtained through importance decomposition, as follows:

$$\mathcal{L}^{\text{ELBO}}(\theta, \vartheta) = \mathbb{E}_{q_\vartheta(X|Y)} \left[\log \frac{p_\theta(Y|X)p_\theta(X)}{q_\vartheta(X|Y)} \right] \quad (1)$$

The latent approximate posterior distribution $q_\vartheta(X|Y)$, the latent prior distribution $p_\theta(X)$, and the measurement likelihood $p_\theta(Y|X)$ of the ELBO (1) are parameterized, respectively, by the neural network (NN) models: encoder, prior, decoder, known collectively as VAE. The parameters θ and ϑ represent the NN parameters that govern these variational models.

To capture the temporal characteristics in a given sequence $y_{\leq L}$ of measurement $y \in \mathbb{R}^{n_y}$ with sequence length L , we similarly introduce a sequence $x_{\leq L}$ of latent states $x \in \mathbb{R}^{n_x}$, and then apply ancestral factorization to the generative and (approximate) posterior distributions, respectively, to obtain

$$\begin{aligned} p_\theta(y_{\leq L}, x_{\leq L}) &= \prod_{n=0}^L p_\theta(y_n | x_{\leq n}, y_{< n}) p_\theta(x_n | x_{< n}, y_{< n}) \\ q_\vartheta(x_{\leq L} | y_{\leq L}) &= \prod_{n=0}^L q_\vartheta(x_n | y_{\leq n}, x_{< n}) \end{aligned} \quad (2)$$

where n denotes the sample index, and $y_{\leq n}, x_{\leq n}$ denote the partial state and measurement sequences up to n^{th} samples.

Let $Y = y_{\leq L}, X = x_{\leq L}$ and substitute (2) into ELBO (1), we obtain the following sequential ELBO (sELBO):

$$\begin{aligned} \mathcal{L}^{\text{sELBO}}(\theta, \vartheta) &= \mathbb{E}_{q_\vartheta(x_{\leq n} | y_{\leq n})} \left[\sum_{n=0}^L \log \frac{p_\theta(y_n | x_{\leq n}, y_{< n}) p_\theta(x_n | x_{< n}, y_{< n})}{q_\vartheta(x_n | y_{\leq n}, x_{< n})} \right] \\ &= \sum_{n=0}^L \mathbb{E}_{q_\vartheta(x_{\leq n} | y_{\leq n})} [\log p_\theta(y_n | x_{\leq n}, y_{< n})] \\ &\quad - \mathcal{D}^{\text{KL}} [q_\vartheta(x_n | y_{\leq n}, x_{< n}) || p_\theta(x_n | x_{< n}, y_{< n})] \end{aligned} \quad (3)$$

where \mathcal{D}^{KL} denotes the (positive-valued) Kullback–Leibler divergence (KLD). The conditional probabilities here encapsulate the underlying causal structure and temporal coherence of sequential data. This sELBO forms the basis of our proposed DSVB framework.

B. Variational Recurrent Neural Network

In this subsection, we introduce the variational RNN (VRNN) parameterization which allows stochastic gradient descent optimization of the VB objective. First, we let the conditional latent prior distribution and measurement likelihood in (2) be the following Gaussian distributions:

$$p_\theta(y_n | x_{\leq n}, y_{< n}) = \mathcal{N}(\mu_n^y, \Sigma_n^y) \quad (5a)$$

$$p_\theta(x_n | x_{< n}, y_{< n}) = \mathcal{N}(\mu_n^x, \Sigma_n^x) \quad (5b)$$

and let the conditional (approximate) posterior distribution in (2) be the following Gaussian distribution:

$$q_\vartheta(x_n | x_{\leq n}, y_{< n}) = \mathcal{N}(\mu_n^{xy}, \Sigma_n^{xy}) \quad (6)$$

with isotropic covariances $\Sigma_n^y = \text{Diag}(\sigma_n^{y^2})$, $\Sigma_n^x = \text{Diag}(\sigma_n^{x^2})$, $\Sigma_n^{xy} = \text{Diag}(\sigma_n^{xy^2})$, where $\text{Diag}(\cdot)$ denotes the diagonal function.

To allow a stochastic gradient decent optimization of the sELBO (3), we then apply the VRNN [35] which parameterizes the mean, standard deviation pairs: (μ_n^y, σ_n^y) , (μ_n^x, σ_n^x) of (5) and $(\mu_n^{xy}, \sigma_n^{xy})$ of (6) as

$$[\mu_n^y, \sigma_n^y] = \varphi_\theta^{\text{dec}}(\varphi_\theta^x(x_n), h_n) \quad (7a)$$

$$[\mu_n^x, \sigma_n^x] = \varphi_\theta^{\text{prior}}(h_n) \quad (7b)$$

$$[\mu_n^{xy}, \sigma_n^{xy}] = \varphi_\vartheta^{\text{enc}}(\varphi_\theta^y(y_n), h_n) \quad (7c)$$

with the memory encoding hidden state

$$h_n = \varphi_\theta^{\text{rnn}}(\varphi_\theta^y(y_{n-1}), \varphi_\theta^x(x_{n-1}), h_{n-1}). \quad (8)$$

Here, the decoder model $\varphi_\theta^{\text{dec}}$, the prior model $\varphi_\theta^{\text{prior}}$, the encoder model $\varphi_\vartheta^{\text{enc}}$, the measurement encoding model φ_θ^y , and the state encoding model φ_θ^x are modelled by deep NNs (DNNs); the RNN model $\varphi_\theta^{\text{rnn}}$ is modelled by either the Long Short-Term Memory (LSTM) [36] or the Gated Recurrent Units (GRU) [37] cell.

To prevent stochastic gradient with high variance [34], the KLD \mathcal{D}^{KL} of sELBO (3) can be analytically solved to obtain

$$\begin{aligned} & - \mathcal{D}^{\text{KL}} [q_\vartheta(x_n | y_{\leq n}, x_{< n}) || p_\theta(x_n | x_{< n}, y_{< n})] \\ &= \int \mathcal{N}(\mu_n^{xy}, \Sigma_n^{xy}) \log \frac{\mathcal{N}(\mu_n^x, \Sigma_n^x)}{\mathcal{N}(\mu_n^{xy}, \Sigma_n^{xy})} dx \\ &= \frac{1}{2} \sum_{j=1}^{n_x} \left[\log \frac{\sigma_{n,j}^{xy^2}}{\sigma_{n,j}^{x^2}} - \frac{\sigma_{n,j}^{xy^2}}{\sigma_{n,j}^{x^2}} - \frac{(\mu_{n,j}^{xy} - \mu_{n,j}^x)^2}{\sigma_{n,j}^{x^2}} + 1 \right] \end{aligned} \quad (9)$$

where $x_{n,j}$ denotes the j^{th} element of x_n . Notice that the analytical solution of the KLD here is deterministic since it does not rely on Monte Carlo (MC) approximation.

Substituting (9) into the sELBO (3) and then expanding the measurement log-likelihood (5a), we have

$$\begin{aligned} \mathcal{L}^{\text{sELBO}}(\theta, \vartheta) &= \frac{1}{2N} \sum_{n=0}^L \sum_{j=1}^{n_x} \sum_{i=1}^N \left[\log(2\pi\sigma_{n,j}^{y^i^2}) + \frac{(y_n - \mu_{n,j}^{y^i})^2}{\sigma_{n,j}^{y^i^2}} \right. \\ &\quad \left. + \log \frac{\sigma_{n,j}^{xy^2}}{\sigma_{n,j}^{x^2}} - \frac{\sigma_{n,j}^{xy^2}}{\sigma_{n,j}^{x^2}} - \frac{(\mu_{n,j}^{xy} - \mu_{n,j}^x)^2}{\sigma_{n,j}^{x^2}} + 1 \right] \end{aligned} \quad (10)$$

$$\begin{aligned}
\mathcal{L}^{\text{DSVB}}(\theta, \vartheta, \phi) &= \mathcal{L}^{\text{sELBO}}(\theta, \vartheta) + \mathcal{L}^{\text{SS}}(\theta, \vartheta) + \mathcal{L}^{\text{BCE}}(\phi) \\
&= \frac{1}{2N} \sum_{n=0}^L \sum_{j=1}^{n_x} \sum_{i=1}^N \left[\overbrace{\log(2\pi\sigma_{n,j}^{y^i}) + \frac{(y_n - \mu_{n,j}^{y^i})^2}{\sigma_{n,j}^{y^i}} + \log \frac{\sigma_{n,j}^{xy^i}}{\sigma_{n,j}^{x^i}} - \frac{\sigma_{n,j}^{xy^i}}{\sigma_{n,j}^{x^i}} - \frac{(\mu_{n,j}^{xy^i} - \mu_{n,j}^{x^i})^2}{\sigma_{n,j}^{x^i}} + 1}_{\mathcal{L}^{\text{sELBO}}(\theta, \vartheta)} \right. \\
&\quad \left. + \mathbb{I}(x_n^* \text{ available}) \left(\underbrace{\log(2\pi\sigma_{n,j}^{x^i}) + \frac{(x_n^* - \mu_{n,j}^{x^i})^2}{\sigma_{n,j}^{x^i}}}_{\mathcal{L}^{\text{SS}}(\theta, \vartheta)} \right) - \underbrace{\left(a^i \log \varphi_\phi^{\text{disc}}(x_n^i) + (1 - a^i) \log(1 - \varphi_\phi^{\text{disc}}(x_n^i)) \right)}_{\mathcal{L}^{\text{BCE}}(\phi)} \right] \quad (4)
\end{aligned}$$

where i denotes the particle index, N denotes the number of particles. Here, we have used the MC approximation $\mathbb{E}_{q(x)}[f(x)] = \frac{1}{N} \sum_{i=1}^N f(x^i)$ with sampled particles $x^i \sim q(x)$ for the measurement log-likelihood expectation. The $(\mu_n^{y^i}, \sigma_n^{y^i})$ is computed via the decoder model (7a) using latent state particles $x_n^i \sim q_\vartheta(x_n | x_{\leq n}, y_{< n})$, obtained using the reparameterization trick $x_n^i = \mu_{n,j}^{xy^i} + \epsilon_{n,j}^i \odot \sigma_{n,j}^{xy^i}$ where $\epsilon_{n,j}^i \sim \mathcal{N}(0, 1)$ and $(\mu_n^{xy^i}, \sigma_n^{xy^i})$ is computed via the encoder model (7c), and \odot denotes the Hadamard product.

C. Disentangled Representation Learning

In general, VRNNs rely on internal latent feature representations that do not have concrete semantics nor tangible interpretations [38], [39]. One way to impose an interpretable and disentangled latent state representation is via the supervision of partial state labels. In this section, we incorporate a semi-supervised disentangled representation learning scheme to endow the latent space of our proposed DSVB with actual system state representation.

Considering the difficulty in acquiring state labels for soft robots, this proposed scheme supervises the conditional latent prior (5b) using partial state label x^* , collected only on certain robot configurations (source domain). This state supervision on the source domain precisely specifies the axes of variation and promotes a disentangled latent representation that concurs with the actual system state representation.

In particular, we incorporate an additional log-likelihood $\mathcal{L}^{\text{SS}}(\theta, \vartheta)$ to sELBO (10) which gives the following semi-supervised sELBO:

$$\begin{aligned}
\mathcal{L}_{\text{SS}}^{\text{sELBO}}(\theta, \vartheta) &= \mathcal{L}^{\text{sELBO}}(\theta, \vartheta) + \mathbb{I}(x_n^* \text{ available}) \mathcal{L}^{\text{SS}}(\theta, \vartheta) \\
\mathcal{L}^{\text{SS}}(\theta, \vartheta) &= \sum_{n=0}^L \mathbb{E}_{q_\vartheta(x_{\leq n} | y_{\leq n})} [\log p_\theta(x_n^* | x_{< n}, y_{< n})] \\
&= \frac{1}{2N} \sum_{n=0}^L \sum_{j=1}^{n_x} \sum_{i=1}^N \left[\log(2\pi\sigma_{n,j}^{x^i}) + \frac{(x_n^* - \mu_{n,j}^{x^i})^2}{\sigma_{n,j}^{x^i}} \right] \quad (11)
\end{aligned}$$

where $\mathbb{I}(\cdot)$ denotes the indicator function. Similarly, here we have expanded the latent prior (5b) in $\mathcal{L}^{\text{SS}}(\theta, \vartheta)$ and used the MC approximation for its expectation. The $(\mu_n^{x^i}, \sigma_n^{x^i})$ is computed via the prior model (7b) using the hidden state particles h_n^i , output by the RNN model (8).

Based on this semi-supervised sELBO (11), the latent prior (5b) is partially conditioned on state labels x_n^* , thus endowing it with the disentangled state representation of the actual system. The latent (approximate) posterior (6)

then inherits this state representation via the KLD (9), which minimizes the distance between the latent prior and (approximate) posterior.

D. Probabilistic Domain Adversarial Training

In practice, it is common for complex soft robot systems under different robot configurations to exhibit distinct dynamics and data distributions. Additionally, the disentangled representation learning described in the previous section can only be applied to the source domain, where state labels are available. However, the lack of state labels in the target domain can lead to an unconditioned latent space, potentially resulting in entangled representations that do not accurately reflect the actual system state representation. This limitation can negatively impact state inference performance.

Taking these into account, in this subsection, we incorporate a probabilistic latent space transfer strategy to encourage domain-adaptable latent features. This transfer strategy facilitate the adaptation of the disentangled representation constructed for the supervised source domain (using partial state labels), to the latent space on the unsupervised target domain. Thus, our objective is to minimize the discrepancy between the latent distributions of the source and target domains. To achieve this, we introduce a probabilistic binary cross entropy (BCE) loss, given by:

$$\begin{aligned}
\mathcal{L}^{\text{BCE}}(\phi) &= \frac{1}{N} \sum_{i=1}^N - \left(a^i \log \varphi_\phi^{\text{disc}}(x_n^i) \right. \\
&\quad \left. + (1 - a^i) \log(1 - \varphi_\phi^{\text{disc}}(x_n^i)) \right) \quad (12)
\end{aligned}$$

where $a^i \in \{0, 1\}$ is the domain label which indicates if the latent sample x_n^i belongs to source or target domain, and $\varphi_\phi^{\text{disc}}$ is a NN-parameterized discriminator model that predicts the domain label given the latent sample. Incorporating (12) to the semi-supervised ELBO (11), we obtain the final DSVB loss $\mathcal{L}^{\text{DSVB}}$ as equation (4).

Motivated by the deterministic domain adversarial training (DAT) strategy [40], we optimize the parameters $(\theta, \vartheta, \phi)$ in an adversarial fashion as follows:

$$\begin{aligned}
(\theta, \vartheta) &= \arg \max_{\theta, \vartheta} \mathcal{L}^{\text{DSVB}}(\theta, \vartheta, \phi) \\
\phi &= \arg \min_{\phi} \mathcal{L}^{\text{DSVB}}(\theta, \vartheta, \phi) \quad (13)
\end{aligned}$$

Under this training strategy, the VRNN parameters (θ, ϑ) are optimized to generate latent samples z^i that fool the classifier $\varphi_\phi^{\text{disc}}$, making the corresponding domain label a^i unrecognizable. At the same time, the discriminator model parameter

ϕ is optimized to accurately distinguish the domain label of the generated latent sample. Such an adversarial competition is expected to achieve the Nash equilibrium that generalizes the latent space to both the source and target domains [32].

III. CASE STUDY

A. Soft Robot System Description

In this work, we adopt a pneumatic-based soft finger (PSF) as shown in Fig. 1A to be used in our case study. This PSF comprises a popular class of soft actuators known as PneuNet [41], commonly used in a soft gripper to pick up odd-shaped and fragile objects which is challenging for conventional rigid grippers [42]. The PSF consists of two sections: the main body with a series of pleated pneumatic chambers and an in-extensible base layer embedded with a resistive flex sensor (Spectra Symbol).

During the experiment, the PSF is actuated using a series of pseudo-random pressure levels within a predetermined pressure range. Two different actuation patterns, namely, *Oscillatory Actuation* and *Random Actuation* are used to verify the generality of our approach to different input signals. In the former, the PSF is actuated using a gradual oscillatory input pressure, while the latter uses a faster random pressure actuation pattern. For each actuation pattern, two experimental configurations were captured. In the first configuration, *Tip Contact*, the PSF is actuated and exposed to a contact bulb placed close to its tip, as shown in Fig. 1A (Left). In the second configuration, *Surface Contact*, the PSF is configured to randomly move along the Z -axis with its entire front surface exposed to the contact bulb, as shown in Fig. 1A (Right). These configurations simulate object grasping at arbitrary locations along the PSF.

During data collection, 10 reflective camera markers are placed evenly along the inextensible base layer to capture the PSF motion. The marker positions are recorded by three motion cameras (OptiTrack Flex13, NaturalPoint Inc.) as 2D Cartesian coordinates, as shown in Fig. 2. The contact forces are measured along axes X, Z using a load cell (Axia80, ATI Industrial Automation Inc.) attached to the contact bulb, as depicted in Fig. 2. Data is sampled at 10Hz, and each sample includes the actuation pressure, flex sensor reading, marker positions and contact forces. On each actuation pattern and experimental configuration, we sample 15000 data points for training and 3000 data points for testing.

B. Experimental Setup

To assess the domain-adaptability of the developed DSVB framework under missing state labels, we supervise the latent state using only state labels from the source domain. The performance of the DSVB is then evaluated through state inference on data from the target domain only. In this work, we design four transfer scenarios with different source and target domains composed of permutations of the captured experimental configurations and actuation patterns. Concretely, these are *Tip Contact* as the source domain with *Surface Contact* as the target domain (Scenario 1) and vice versa (Scenario 2). Alternatively, we also have *Oscillatory*

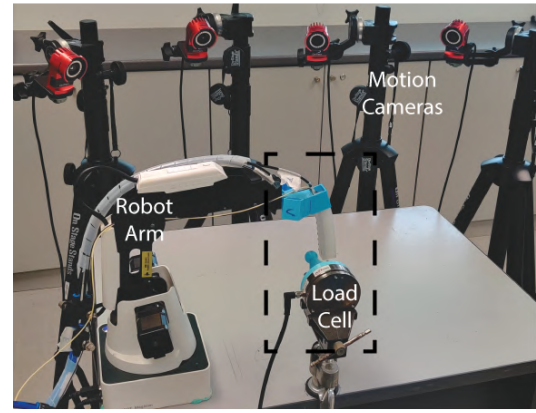


Fig. 2. **The setup of pneumatic-based soft finger (PSF).** Four motion cameras are used to track the markers attached on the PSF. A rigid robot arm is used to maneuver the position of PSF.

Actuation as the source domain with *Random Actuation* as the target domain (Scenario 3) and vice versa (Scenario 4). By training and testing on differing source and target domains, we demonstrate the ability of the proposed method in transfer learning the well learned representations from the supervised source domain to the unseen target domain. We benchmark the proposed DSVB against two state-of-the-art RNNs, Gated Recurrent Unit (GRU) and Long Short-Term Memory (LSTM) networks. In contrast to the proposed method, these models are deterministic and are trained using only the partial measurement data and state label pairs of the source domain.

In this case study, we validate two implementations of the DSVB, where the RNN model φ_{θ}^{rnn} is either parameterized by a single-layer vanilla GRU [37], namely DSVB-GRU, or a single-layer LSTM [36], DSVB-LSTM. In both cases, the RNN model uses a hidden state $h_n \in \mathbb{R}^{128}$ of size 128. The prior model φ_{θ}^{prior} , encoder model φ_{θ}^{enc} and the decoder model φ_{θ}^{dec} are feed-forward NNs (FNNs) with hidden layer architectures $\{128\}$, $\{128, 128\}$ and $\{128, 64\}$, respectively, where each entry of the curly brackets is a hidden layer and the value denotes layer width. The state and measurement encoding models φ_{θ}^x and φ_{θ}^y , respectively, are FNNs with hidden layer architectures $\{32\}$ and $\{22\}$. The domain classifier φ_{ϕ}^{disc} is parameterized by a single GRU layer with 128 hidden states and a succeeding FNN with hidden layer architecture $\{128, 1\}$. For the benchmark methods, we consider a vanilla GRU or LSTM with 128 hidden states, followed by a FNN $\{128\}$ that maps the hidden states to the system states. The NN parameters across all these comparison methods are initiated with the same values.

All the comparison methods (DSVB and benchmark methods) are trained on a sequence length of 100 and a batch size of 128, using the Adam optimizer with a learning rate of 0.001 for 50 epochs to allow the methods to converge. During inference time, the comparison methods output the estimated latent state labels (contact forces and marker positions) using the original measurement data (flex reading and pressure input) from the test set. The latent state labels in the test set serve as ground truth for result validation.

TABLE I
COMPARISON BETWEEN STATE INFERENCE RESULTS OF THE PROPOSED DSVB AND BENCHMARK STATE-OF-THE-ART RECURRENT NEURAL NETWORKS ON SOURCE AND TARGET DOMAINS

Method	Scenario 1		Scenario 2		Scenario 3		Scenario 4	
	Source (Tip Contact)	Target (Surface Contact)	Source (Surface Contact)	Target (Tip Contact)	Source (Osc. Actuation)	Target (Rand. Actuation)	Source (Rand. Actuation)	Target (Osc. Actuation)
LSTM	0.952±0.024	2.678±0.074	0.448±0.014	1.087±0.026	0.939±0.032	1.564±0.028	0.883±0.009	1.522±0.022
DSVB-LSTM	1.237±0.053	0.961±0.017	0.594±0.034	0.680±0.049	1.187±0.021	1.005±0.014	0.925±0.046	1.058±0.038
GRU	0.939±0.024	2.863±0.065	0.442±0.013	1.096±0.012	0.915±0.031	1.577±0.040	0.890±0.022	1.487±0.020
DSVB-GRU	1.179±0.009	0.940±0.025	0.551±0.017	0.673±0.049	1.157±0.069	0.986±0.017	0.886±0.026	1.021±0.044

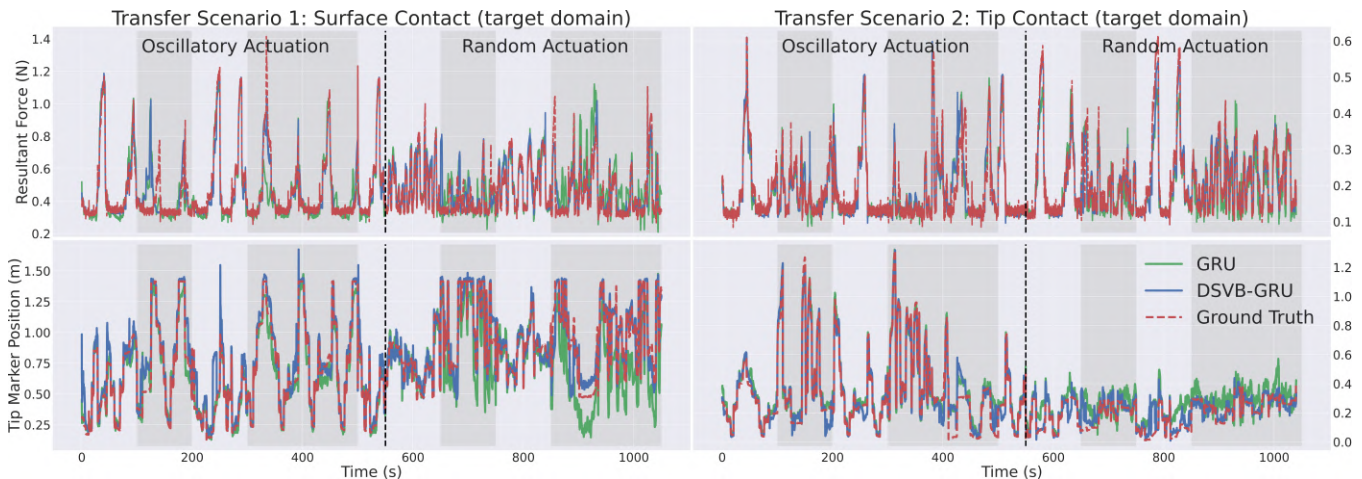


Fig. 3. Normalized time-series state estimation of the best benchmark (GRU) and proposed method (DSVB-GRU) on the target domains of transfer Scenario 1 (left) and Scenario 2 (right). Resultant force (top) is the 2-norm of contact forces X, Z and tip marker position (bottom) is the 2-norm of tip marker coordinates X, Z . The shaded and unshaded regions indicate time windows with states from the (unsupervised) target domain and (supervised) source domain, respectively.

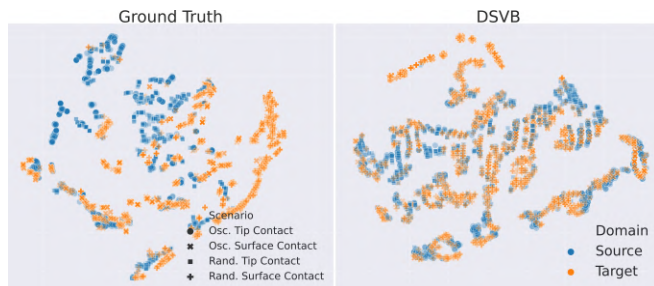


Fig. 4. Comparison of tSNE visualizations of source and target domains for the ground truth and the posterior latent states in the proposed DSVB.

C. Results & Discussions

Table I presents the prediction performance (root-mean-squared errors) of the four scenarios described, where the means and standard deviations are aggregated over five runs. For an ablative study, we compare the RNN benchmark models (either GRU or LSTM) directly with the corresponding DSVB implementations. The results show that our proposed DSVB framework achieves comparable performances to the benchmark methods when evaluated on the supervised source domain. However, the differences in state estimation error between the proposed DSVB and the benchmark methods become much more significant on the unsupervised target domain, which contains system dynamics that were not directly exposed to the models during training. Figure 3 illustrates that the GRU has difficulty producing accurate estimates on the target domains of transfer scenarios 1 and 2, presumably due to the distinct dynamics in the source

and target domains. In contrast, the DSVB yields consistent results across both domains, indicating that the proposed framework achieves effective transfer learning.

Furthermore, to verify the capability of the probabilistic domain adversarial training scheme, Figure 4 shows the tSNE distribution for the source and target domains from the ground truth and the proposed DSVB method for Scenario 1. Here, we observe that the tSNE embeddings of the proposed method exhibit smaller discrepancies between the two domains compared to the ground truth. The affinity of the source and target domains highlights the capability of the DSVB framework in transferring the disentangled representations acquired from the supervised source domain to achieve domain-adaptable latent features which facilitate cross-domain transfer learning even in the absence of missing state labels on the target domain.

IV. CONCLUSION

The capability of DSVB to perform transfer learning offers a solution for accurate soft robot modelling and inference in the absence of adequate state labels. The experimental results show that labelled data only need to be collected on relatively simple source robot configurations, and the knowledge learnt can be transferred by the proposed framework to multiple unlabelled target configurations. For future works, we could investigate the feasibility of DSVB on more involved non-planar soft robots [43]–[45]. We plan to also incorporate normalizing flows [46], [47] to our proposed DSVB.

REFERENCES

- [1] C. Lee, M. Kim, Y. J. Kim, N. Hong, S. Ryu, H. J. Kim, and S. Kim, "Soft robot review," *International Journal of Control, Automation and Systems*, vol. 15, no. 1, pp. 3–15, 2017.
- [2] C. Laschi, B. Mazzolai, and M. Cianchetti, "Soft robotics: Technologies and systems pushing the boundaries of robot abilities," *Science Robotics*, vol. 1, no. 1, p. eaah3690, 2016.
- [3] S. Kim, C. Laschi, and B. Trimmer, "Soft robotics: a bioinspired evolution in robotics," *Trends in biotechnology*, vol. 31, no. 5, pp. 287–294, 2013.
- [4] S. G. Nurzaman, F. Iida, C. Laschi, A. Ishiguro, and R. Wood, "Soft robotics [tc spotlight]," *IEEE Robotics & Automation Magazine*, vol. 20, no. 3, pp. 24–95, 2013.
- [5] K. Elgeneidy, N. Lohse, and M. Jackson, "Bending angle prediction and control of soft pneumatic actuators with embedded flex sensors—a data-driven approach," *Mechatronics*, vol. 50, pp. 234–247, 2018.
- [6] T. G. Thuruthel, B. Shih, C. Laschi, and M. T. Tolley, "Soft robot perception using embedded soft sensors and recurrent neural networks," *Science Robotics*, vol. 4, no. 26, 2019.
- [7] R. Wang, S. Wang, S. Du, E. Xiao, W. Yuan, and C. Feng, "Real-time soft body 3d proprioception via deep vision-based sensing," *IEEE Robot. Autom. Lett.*, vol. 5, no. 2, pp. 3382–3389, 2020.
- [8] J. Y. Loo, Z. Y. Ding, V. M. Baskaran, S. G. Nurzaman, and C. P. Tan, "Robust multimodal indirect sensing for soft robots via neural network-aided filter-based estimation," *Soft Robotics*, 2021.
- [9] D. Kim, S.-H. Kim, T. Kim, B. B. Kang, M. Lee, W. Park, S. Ku, D. Kim, J. Kwon, H. Lee *et al.*, "Review of machine learning methods in soft robotics," *PLoS One*, vol. 16, no. 2, p. e0246102, 2021.
- [10] K. Chin, T. Hellebrekers, and C. Majidi, "Machine learning for soft robotic sensing and control," *Advanced Intelligent Systems*, p. 1900171, 2020.
- [11] M. Cianchetti, M. Calisti, L. Margheri, M. Kuba, and C. Laschi, "Bioinspired locomotion and grasping in water: the soft eight-arm octopus robot," *Bioinspiration & biomimetics*, vol. 10, no. 3, p. 035003, 2015.
- [12] A. D. Marchese, C. D. Onal, and D. Rus, "Autonomous soft robotic fish capable of escape maneuvers using fluidic elastomer actuators," *Soft Robotics*, vol. 1, no. 1, pp. 75–87, 2014.
- [13] A. Khan, W. Liquan, W. Gang, M. Imran, H. M. Waqas, and A. A. Zaidi, "Concept design of the underwater manned seabed walking robot," *Journal of Marine Science and Engineering*, vol. 7, no. 10, p. 366, 2019.
- [14] K. J. Cios and G. W. Moore, "Uniqueness of medical data mining," *Artificial intelligence in medicine*, vol. 26, no. 1-2, pp. 1–24, 2002.
- [15] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, and Q. He, "A comprehensive survey on transfer learning," *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43–76, 2021.
- [16] H. Wang, M. Totaro, and L. Beccai, "Toward perceptive soft robots: Progress and challenges," *Advanced Science*, vol. 5, no. 9, p. 1800541, 2018.
- [17] G. Fang, X. Wang, K. Wang, K.-H. Lee, J. D. Ho, H.-C. Fu, D. K. C. Fu, and K.-W. Kwok, "Vision-based online learning kinematic control for soft robots using local gaussian process regression," *IEEE Robot. Autom. Lett.*, vol. 4, no. 2, pp. 1194–1201, 2019.
- [18] G. Runge, M. Wiese, and A. Raatz, "Fem-based training of artificial neural networks for modular soft robots," in *2017 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE, 2017, pp. 385–392.
- [19] L. Massari, E. Schena, C. Massaroni, P. Saccomandi, A. Menciasci, E. Sinibaldi, and C. M. Oddo, "A machine-learning-based approach to solve both contact location and force in soft material tactile sensors," *Soft robotics*, vol. 7, no. 4, pp. 409–420, 2020.
- [20] G. Runge, M. Wiese, and A. Raatz, "Fem-based training of artificial neural networks for modular soft robots," in *2017 ROBIO*. IEEE, 2017, pp. 385–392.
- [21] H. Sui, W. Shang, and X. Li, "Transfer of robot perception module with adversarial learning," *IEEE Access*, vol. 7, pp. 79726–79736, 2019.
- [22] S. Kriegman, A. M. Nasab, D. Shah, H. Steele, G. Branin, M. Levin, J. Bongard, and R. Kramer-Bottiglio, "Scalable sim-to-real transfer of soft robot designs," in *2020 3rd IEEE International Conference on Soft Robotics (RoboSoft)*, 2020, pp. 359–366.
- [23] H. Park, J. Cho, J. Park, Y. Na, and J. Kim, "Sim-to-real transfer learning approach for tracking multi-dof ankle motions using soft strain sensors," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 3525–3532, 2020.
- [24] H. Donat, S. Lilge, J. Burgner-Kahrs, and J. J. Steil, "Estimating tip contact forces for concentric tube continuum robots based on backbone deflection," *IEEE Transactions on Medical Robotics and Bionics*, vol. 2, no. 4, pp. 619–630, 2020.
- [25] M. Wiese, G. Runge-Borchert, B.-H. Cao, and A. Raatz, "Transfer learning for accurate modeling and control of soft actuators," in *2021 IEEE 4th International Conference on Soft Robotics (RoboSoft)*, 2021, pp. 51–57.
- [26] G. Fang, Y. Tian, Z.-X. Yang, J. M. P. Geraedts, and C. C. L. Wang, "Efficient jacobian-based inverse kinematics with sim-to-real transfer of soft robots by learning," *IEEE/ASME Transactions on Mechatronics*, pp. 1–11, 2022.
- [27] W. Jiang, H. Gao, W. Lu, W. Liu, F.-L. Chung, and H. Huang, "Stacked robust adaptively regularized auto-regressions for domain adaptation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 31, no. 3, pp. 561–574, 2019.
- [28] S. Yang, H. Wang, Y. Zhang, P. Li, Y. Zhu, and X. Hu, "Semi-supervised representation learning via dual autoencoders for domain adaptation," *Knowledge-Based Systems*, vol. 190, p. 105161, 2020.
- [29] S. Yang, K. Yu, F. Cao, H. Wang, and X. Wu, "Dual-representation-based autoencoder for domain adaptation," *IEEE Transactions on Cybernetics*, pp. 1–14, 2021.
- [30] F. Zhuang, X. Cheng, P. Luo, S. J. Pan, and Q. He, "Supervised representation learning with double encoding-layer autoencoder for transfer learning," *ACM Trans. Intell. Syst. Technol.*, vol. 9, no. 2, oct 2017.
- [31] L. Wen, L. Gao, and X. Li, "A new deep transfer learning based on sparse auto-encoder for fault diagnosis," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 1, pp. 136–144, 2019.
- [32] Z. Chai, C. Zhao, B. Huang, and H. Chen, "A deep probabilistic transfer learning framework for soft sensor modeling with missing data," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–12, 2021.
- [33] B. Schölkopf, F. Locatello, S. Bauer, N. R. Ke, N. Kalchbrenner, A. Goyal, and Y. Bengio, "Toward causal representation learning," *Proceedings of the IEEE*, vol. 109, no. 5, pp. 612–634, 2021.
- [34] D. Kingma and M. Welling, "Auto-encoding variational bayes," in *The 2nd International Conference on Learning Representations (ICLR)*, 2013.
- [35] J. Chung, K. Kastner, L. Dinh, K. Goel, A. C. Courville, and Y. Bengio, "A recurrent latent variable model for sequential data," in *Advances in Neural Information Processing Systems*, vol. 28. Curran Associates, Inc., 2015.
- [36] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [37] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," *arXiv preprint arXiv:1406.1078*, 2014.
- [38] S. N. B. Paige, J.-W. van de Meent, A. Desmaison, N. Goodman, P. Kohli, F. Wood, and P. Torr, "Learning disentangled representations with semi-supervised deep generative models," in *Advances in Neural Information Processing Systems*, vol. 30. Curran Associates, Inc., 2017.
- [39] D. P. Kingma, S. Mohamed, D. Jimenez Rezende, and M. Welling, "Semi-supervised learning with deep generative models," in *Advances in Neural Information Processing Systems*, vol. 27. Curran Associates, Inc., 2014.
- [40] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, and V. Lempitsky, "Domain-adversarial training of neural networks," *The journal of machine learning research*, vol. 17, no. 1, pp. 2096–2030, 2016.
- [41] F. Ilievski, A. D. Mazzeo, R. F. Shepherd, X. Chen, and G. M. Whitesides, "Soft robotics for chemists," *Angewandte Chemie*, vol. 123, no. 8, pp. 1930–1935, 2011.
- [42] S. G. Nurzaman, F. Iida, L. Margheri, and C. Laschi, "Soft robotics on the move: Scientific networks, activities, and future challenges," *Soft Robot.*, vol. 1, no. 2, pp. 154–158, 2014.
- [43] K. Nakajima, H. Hauser, T. Li, and R. Pfeifer, "Exploiting the

- dynamics of soft materials for machine learning,” *Soft Robotics*, vol. 5, no. 3, pp. 339–347, 2018.
- [44] D. Bruder, X. Fu, R. B. Gillespie, C. D. Remy, and R. Vasudevan, “Data-driven control of soft robots using koopman operator theory,” *IEEE Transactions on Robotics*, pp. 1–14, 2020.
- [45] I. M. Van Meerbeek, C. M. De Sa, and R. F. Shepherd, “Soft optoelectronic sensory foams with proprioception,” *Science Robotics*, vol. 3, no. 24, 2018.
- [46] D. Rezende and S. Mohamed, “Variational inference with normalizing flows,” in *Proceedings of the 32nd International Conference on Machine Learning*, vol. 37. PMLR, 2015, pp. 1530–1538.
- [47] D. P. Kingma, T. Salimans, R. Jozefowicz, X. Chen, I. Sutskever, and M. Welling, “Improved variational inference with inverse autoregressive flow,” in *Advances in Neural Information Processing Systems*, vol. 29. Curran Associates, Inc., 2016.