

# SEGMENTATION OF THE EVOLVING LEFT VENTRICLE BY LEARNING THE DYNAMICS

Walter Sun<sup>a</sup>, Müjdat Çetin<sup>b,c</sup>, Ray Chan<sup>d</sup>, and Alan S. Willsky<sup>c</sup>

<sup>a</sup>Microsoft Corporation, Redmond, WA, USA

<sup>b</sup>Faculty of Engineering and Natural Sciences, Sabancı University, İstanbul, Turkey

<sup>c</sup>Laboratory for Information and Decision Systems, MIT, Cambridge, MA, USA

<sup>d</sup>Massachusetts General Hospital, 55 Fruit Street, Boston, MA, USA

## ABSTRACT

We propose a method for recursive segmentation of the left ventricle (LV) across a temporal sequence of magnetic resonance (MR) images. The approach involves a technique for learning the LV boundary dynamics together with a particle-based inference algorithm on a loopy graphical model capturing the temporal periodicity of the heart. The dynamic system state is a low-dimensional representation of the boundary, and boundary estimation involves incorporating curve evolution into state estimation. By formulating the problem as one of state estimation, the segmentation at each particular time is based not only on the data observed at that instant, but also on predictions based on past and future boundary estimates. We assess and demonstrate the effectiveness of the proposed framework on a large data set of breath-hold cardiac MR image sequences.

**Index Terms**— Magnetic resonance imaging, cardiac imaging, left ventricle, image segmentation, recursive estimation, curve evolution, graphical models, particle filtering, learning.

## 1. INTRODUCTION

Segmentation of the left ventricle (LV) of the heart across a cardiac cycle is a problem of interest because the left ventricle’s proper function, pumping oxygenated blood to the entire body, is vital for normal human activity. Having segmentations of the LV over time allows cardiologists to assess the dynamic behavior of the human heart (using, e.g., the ejection fraction). Observing how the LV evolves throughout an entire cardiac cycle allows physicians to determine the health of the myocardial muscles. Segmented LV boundaries can also be useful for further quantitative analysis. For example, past work [1, 2] on extracting the flow fields of the myocardial wall assumes the availability of LV segmentations throughout the cardiac cycle.

The automated segmentation of the LV endocardium in bright blood cardiac magnetic resonance (MR) images is a challenging problem. Exploiting segmentations from a previous or future frame can improve the segmentation of the LV in the current frame, because the LV boundaries exhibit strong temporal correlation. Using such information would be particularly useful for low SNR images, in which the observation from a single frame alone may not provide sufficient information for an acceptable segmentation. The incorporation of dynamic models into cardiac segmentation and tracking is

an area of recent and growing interest. Chalana, *et al.* [3] and Jolly *et al.* [4] perform causal processing using the segmentation from the most recent frame to initiate a local search for the segmentation in the current frame. Senegas, *et al.* [5] use sample-based methods (using sequential Monte Carlo) for causal shape tracking on a finite-dimensional representation of the shape space using spherical harmonics.

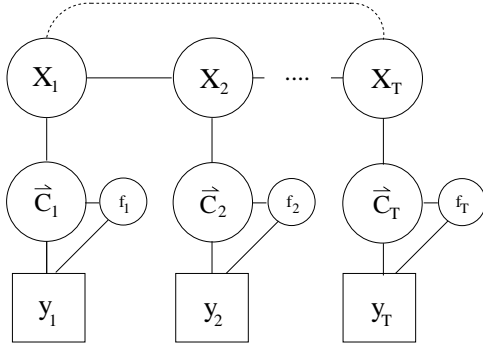
We propose a segmentation approach in which we use level set representations in concert with finite-dimensional projections of curves onto principal components. We use nonparametric, information-theoretic methods to learn statistical models for temporal evolution of these finite-dimensional projections. We use sample-based methods [11] for fusing our dynamic model with the MR image sequence data. We exploit the periodicity of cardiac motion, and perform inference on a graph consisting of a single cycle. While the core of our approach involves sampling and inference on finite-dimensional projections of the shape space, by coupling level set representations into our graphical structure appropriately, our ultimate boundary estimates are in fact curves generated via curve evolution, and hence are not constrained to a finite-dimensional space.

In summary, the major contributions of our approach are (a) the use of nonparametric, information-theoretic methods to learn the dynamics of the LV evolution; (b) the development of a particle-based inference algorithm that allows segmentation throughout the LV cycle using all available data; (c) the incorporation of level set methods in a statistically consistent manner in order to generate boundaries not constrained to reside in the finite-dimensional domain in which dynamics have been learned; and (d) the demonstration of the effectiveness of our approach (both qualitatively and quantitatively) on cardiac MR imagery.

## 2. PROBABILISTIC MODEL STRUCTURE

We formulate the dynamic segmentation problem as the estimation of the posterior distribution of the left ventricular boundary at each discrete time  $t \in \{1, \dots, T\}$ , where  $T$  denotes the number of temporal MR image frames over a single cardiac cycle. Figure 1 is a graphical model representation of the statistical structure we propose for describing data and cardiac dynamics. Here, circular and rectangular nodes correspond to hidden and observed random variables, respectively. At each time point  $t$ , we observe the image data  $y_t$ , which are noisy measurements of the blood and tissue intensity field. Let  $f_t$  be a coarse description of the intensities inside and outside the LV (in particular, while other, more complex descriptions can be used, we take  $f_t$  to be the 2D vector consisting of the mean intensity inside the LV and the mean intensity outside). Let  $\vec{C}_t$  denote the LV boundary

This work was supported in part by a MURI funded through AFOSR Grant FA9550-06-1-0324, in part by Shell International Exploration and Production, Inc., and in part by the European Commission under Grant MIRG-CT-2006-041919.



**Fig. 1.** Graphical model representing the statistical model structure used in our work.

at time  $t$ . One key aspect of this model is that we also define a variable  $X_t$ , which is a low-dimensional representation of  $\vec{C}_t$ , and which we use as the dynamic system state. A definition of  $X_t$  will follow in Section 3. The structure in Figure 1 implies a Markov structure for the dynamical model of  $X_t$ . The dashed line connecting  $X_1$  and  $X_T$  indicates a graphical link to exploit the periodicity of the cardiac cycle, capturing the fact that although  $X_1$  and  $X_T$  are not temporal neighbors in absolute time, they are neighbors in the cardiac cycle.

This model involves a *likelihood model*  $p(y_t|f_t, \vec{C}_t)$ , an *intensity prior model*  $p(f_t|\vec{C}_t)$ , a *curve prior model*  $p(\vec{C}_t|X_t)$ , and a *dynamical model for the system state*  $p(X_t|X_{t-1})$ . Note that each of these models can be associated with a single edge or multiple edges in Figure 1, and also that due to the Markov structure, these are the only relationships required to characterize the joint distribution of the variables involved. We specify a simple curve prior model  $p(\vec{C}_t|X_t)$  that penalizes the deviation of  $\vec{C}_t$  from the curve represented by  $X_t$ . For the intensity prior model  $p(f_t|\vec{C}_t)$ , based on empirical observations, we use a log-normal density, in which the model parameters are learned from intensities of segmented training data. Similarly, for the likelihood model  $p(y_t|f_t, \vec{C}_t)$ , we use a log-normal density whose mean involves the mean intensity variable  $f_t$ . For the exact form of these models, see [6].

The major component needed to complete our statistical model is the Markovian dynamics  $p(X_t|X_{t-1})$  (or  $p(X_{t-1}|X_t)$  in the backward direction) between two temporally-neighboring states  $X_{t-1}$  and  $X_t$  in the evolution of the LV<sup>1</sup>. This can be associated with the edges between the  $X$ 's in Figure 1. The perspective we take in this paper is to learn a time-varying dynamical model from off-line training data (Section 3), and then use that model in the automatic segmentation process (Section 4).

### 3. SYSTEM DYNAMICS

In this section, we describe our finite dimensional representation of LV boundaries, and propose an information-theoretic technique for learning the boundary dynamics.

Using a finite-dimensional approximate representation  $X_t$  of the boundaries  $\vec{C}_t$  makes the learning problem better posed and the learned models statistically significant. We use principal component analysis (PCA) on the signed distance functions (SDFs) of the (area-normalized) boundaries to obtain a finite-dimensional basis (keeping 8 principal components) for the shapes [7, 8]. We append this basis with area to obtain a 9-dimensional basis. Given this basis,

<sup>1</sup>In addition, we have similar models relating  $X_1$  and  $X_T$ .

we can represent the SDF of any given shape by a linear combination of the basis elements. Thus the 9-dimensional system state  $X_t$  is composed of the coefficients for that representation, and we are interested in learning its dynamics.

Unlike existing work (e.g., [9]), we learn the dynamics using information-theoretic ideas and nonparametric statistics, without restrictions of linearity or Gaussianity. Our approach is based on the technique in [10], which provides a nonparametric, yet computationally tractable approach to learning the dynamics. Our approach involves two steps. First we learn maximally informative statistics about the past (or future) for the prediction of the current state. The second step involves non-parametrically estimating the (forward and backward) transition densities.

We are interested in learning the statistical relationship between  $X_{t-1}$  and  $X_t$ . Ideally, we would like to consider only the portion of the state  $X_{t-1}$  that is statistically pertinent to the prediction of  $X_t$ . To this end, we introduce a function  $q_{t-1}(X_{t-1})$  which seeks to reduce dimensionality yet capture all of the information in  $X_{t-1}$  that relates to  $X_t$ <sup>2</sup>. This is achieved exactly only when  $I(X_t; X_{t-1}) = I(X_t; q_{t-1}(X_{t-1}))$ , where  $I(X_t; X_{t-1})$  is the mutual information between  $X_t$  and  $X_{t-1}$ . In practice, even if we cannot generally make  $I(X_t; q_{t-1}(X_{t-1}))$  equal to  $I(X_t; X_{t-1})$ , we can try to make it as large as possible. As a result, we pose the problem as finding  $q_{t-1}$  that maximizes  $I(X_t; q_{t-1}(X_{t-1}))$ . This makes  $q_{t-1}(X_{t-1})$  a maximally-informative statistic. In the experiments presented in this paper, we constrain  $q_t(\forall t)$  to be a linear function. It should be noted that the linearity assumption for  $q_t$  does not mean that the resulting dynamic model is either linear or Gaussian [10]. We then obtain the forward densities (backward densities can be learned in a completely analogous manner) based on training data. We note that the procedure we describe in this section produces densities of the form  $p(X_t|q_{t-1}(X_{t-1}))$  which we use as approximate representations of the state dynamics  $p(X_t|X_{t-1})$ . Given samples of  $X_t$  and  $X_{t-1}$ , we construct a kernel density estimate of the joint density  $p_{X_t, q_{t-1}(X_{t-1})}(X_t, q_{t-1}(X_{t-1}))$ , from which we can obtain the desired forward transition density by conditioning:  $p_{X_t|q_{t-1}(X_{t-1})}(X_t|q_{t-1}(X_{t-1}))$ . Using this framework, one can in principle learn a fully time-varying model. However, in practice, this requires the availability of sufficient training data to support such learning. When the number of training samples is moderate, one could impose a time-invariant model for the systolic phase, and one for the diastolic phase (and hence use all data in that phase, rather than only at a particular time, for learning), as we do in our experiments in Section 5.

### 4. INFERENCE ALGORITHM

In this section we describe our approach to dynamic segmentation based on the probabilistic model introduced in Section 2, and the dynamic model for  $X$  learned using the method discussed in Section 3. Thanks both to the nonparametric form of the forward and backward transition densities for  $X$  and, more generally, the complex nature of the Bayesian inference problems we wish to solve, we are led to the use of sampling - i.e., particle-based algorithms [11] to approximate the densities and likelihood functions required in the inference procedure. However, there are two issues that preclude straightforward use of standard particle filtering methods in our problem. The first is that rather than causal filtering on a chain, we are interested in performing inference on the cycle graph of Figure 1. The second is that a full particle representation of the inference problem of LV segmen-

<sup>2</sup>Note that we can define a function in the "backward" direction in a completely analogous fashion.

tation would involve generating particles for the infinite-dimensional curves,  $\vec{C}$ . We avoid this complexity through an approximation that involves solving level-set based curve evolutions during the sampling procedure.

Our goal is to estimate  $X_t$  for any  $t$  based on  $y_{1:T} = \{y_\tau | 1 \leq \tau \leq T\}$ . The posterior density  $p(X_t | y_{1:T})$  is proportional to:

$$\int_{f_t} \int_{\vec{C}_t} p(y_t | f_t, \vec{C}_t) p(f_t | \vec{C}_t) p(\vec{C}_t | X_t) p(X_t | y_{(1:T) \setminus t}) d\vec{C}_t df_t, \quad (1)$$

where  $y_{(1:T) \setminus t}$  denotes all data in  $y_{1:T}$  except  $y_t$ . One complexity that would be encountered by a particle-based algorithm here is the marginalization operation over  $f_t$  and  $\vec{C}_t$ . We consider the following approximate representation of (1):

$$p(X_t | y_{1:T}) \propto p(y_t | f_t^*, \vec{C}_t^*) p(f_t^* | \vec{C}_t^*) p(\vec{C}_t^* | X_t) p(X_t | y_{(1:T) \setminus t}). \quad (2)$$

We will describe how to compute the new variables  $f_t^*$  and  $\vec{C}_t^*$ , but for the time being, assume we have them. We are then left with the computation of (2) at each point in time. This computation can be performed by a message-passing algorithm called belief propagation (BP) [12]. This algorithm requires the specification of the probabilistic model in terms of *potential functions* corresponding to each node/time in our graph and to each edge. In our case, the node potentials capture all of the statistical information localized to time  $t$  alone, i.e.:

$$\psi_t(X_t, y_t) = p(y_t | f_t^*, \vec{C}_t^*) p(f_t^* | \vec{C}_t^*) p(\vec{C}_t^* | X_t) p(X_t). \quad (3)$$

The edge potentials are given by

$$\psi_{t,\tau}(X_t, X_\tau) = \frac{p(X_t, X_\tau)}{p(X_t)p(X_\tau)} = \frac{p(X_t | X_\tau)}{p(X_t)} = \frac{p(X_\tau | X_t)}{p(X_\tau)}, \quad (4)$$

where the involvement of the system dynamics can be observed. Given these potentials, BP involves an iterative procedure in which every node sends a “message” to its neighbors in each iteration. BP provides an exact solution for tree-structured graphs, and an approximate solution for loopy graphs. The message from node  $\tau$  to node  $t$  at iteration  $n$  is given by:

$$m_{\tau t}^n(X_t) \propto \int_{X_\tau} \psi_{t,\tau}(X_t, X_\tau) \psi_\tau(X_\tau, y_\tau) \prod_{s \in N(\tau) \setminus t} m_{s\tau}^{n-1}(X_\tau) dX_\tau, \quad (5)$$

where  $N(\tau) \setminus t$  denotes all neighbors of node  $X_\tau$  except  $X_t$ . The posterior can be computed as:

$$p^n(X_t | y_{1:T}) \propto \psi_t(X_t, y_t) \prod_{\tau \in N(t)} m_{\tau t}^n(X_t). \quad (6)$$

Such message computations can be performed exactly only in special cases; e.g., when everything is Gaussian. In the more general case involving nonparametric densities, one can perform approximate computations using particle-based methods. In our work, we use the nonparametric belief propagation (NBP) algorithm [13]. In NBP, particle representations serve as weighted mixture distribution approximations (in particular Gaussians sums) of the true messages.

Now, let us turn to an issue we have previously put aside, namely the approximation in (2) involving the variables  $f_t^*$  and  $\vec{C}_t^*$ . To compute these quantities in each message passing iteration  $n$  (with the resulting values then used to update the node potentials in (3)), we carry out the following optimization:

$$(f_t^*[n], \vec{C}_t^*[n]) = \arg \min_{f_t, \vec{C}_t} E_n(f_t, \vec{C}_t) \quad (7)$$

where  $E_n(f_t, \vec{C}_t)$  is given by:

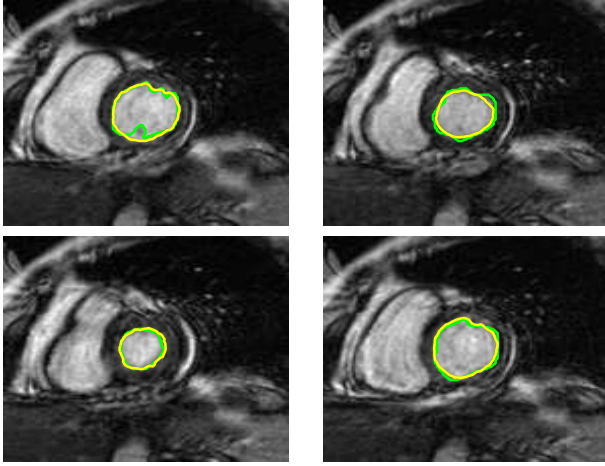
$$-\log p(y_t | f_t, \vec{C}_t) - \log p(f_t | \vec{C}_t) - \log \int_{X_t} p(\vec{C}_t | X_t) \prod_{\tau \in N(t)} m_{\tau t}^n(X_t) dX_t \quad (8)$$

Note that this is an approximate way of replacing the marginalization operation in (1) by the maximization of the integrand [6]. We solve this variational problem by curve evolution based on level sets. Once the values of  $f_t^*[n]$  and  $\vec{C}_t^*[n]$  are found to minimize this functional we update the node potential  $\psi_t^n(X_t, y_t)$  (note that here we indicate the dependence of the potential on the iteration  $n$  explicitly) using (3). Then the NBP belief update procedure [13] can be used to generate samples from the posterior density  $p^n(X_t | y_{1:T})$  in (6). When the NBP iterations converge, we can use the final  $\vec{C}_t^*$  as the segmented LV boundary at time  $t$ .

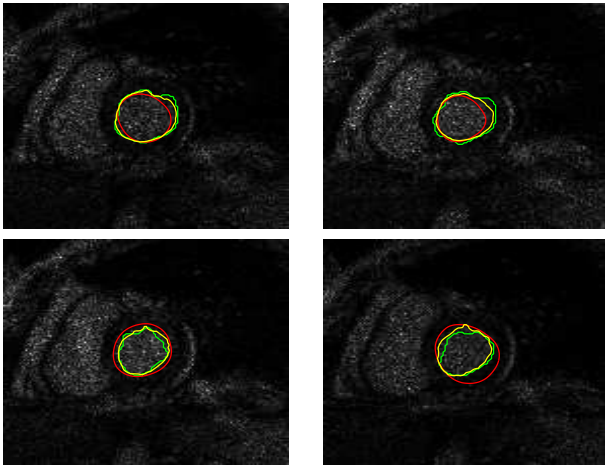
## 5. EXPERIMENTAL RESULTS

We apply the proposed technique on 2-D mid-ventricular slices of MR data, but we also note that we can in principle apply the method to 3-D data. The data set we use contains twenty-frame time sequences of breath-hold cardiac MR images, each representing a single cardiac cycle with the initial frame gated (synchronized) with an electrocardiogram (EKG) signal. We do not consider arrhythmias because only patients having sustained and hemodynamically-stable arrhythmias, a rare situation, can be practically imaged and analyzed. Anonymized data sets were obtained from the Cardiovascular MR-CT Program at the Massachusetts General Hospital. Our training set consists of 42 cardiac cycles of 20 frames each for a total of 840 images, acquired from five patients. The segmentations on the training set were carried out manually by both radiologists as well as researchers whose results were reviewed by radiologists. All of the training and test sets come from healthy patients. Our test set consists of 234 cardiac cycles of 20 frames each, acquired from 26 patients. These patients are different from the ones used for the training set. We measure accuracy of the segmentations by computing the dice coefficient [14], which evaluates to 1 when the segmentation result and the underlying true region perfectly match, and 0 when the regions do not overlap at all. To determine the accuracy of a set of boundary estimates, we compute the dice measure between each of the estimates  $\vec{C}_t^*$  and the ground truth at that time, and then average over all of the dice coefficients computed.

Figure 2 shows our segmentation results on representative frames of high-SNR data, superimposed on the manual segmentations representing the ground truth. The results appear to be quite accurate. In Figure 3, we present some results on low-SNR data. Here, we also display the results of the static, shape prior-based segmentation algorithm of Tsai *et al.* [8] for comparison. The static segmentation method of [8] uses the same training data as our approach for learning the shape priors. We qualitatively observe that our approach leads to more accurate segmentations. This observation is supported quantitatively by the dice boundary coefficients presented in Table 1. The static segmentation method achieves 0.8516 and 0.8035, for the high and low SNR cases respectively, while the corresponding numbers for our algorithm are 0.9292 and 0.9069. This quantitative empirical analysis demonstrates the effectiveness of our strategy in incorporating information from other temporal frames into the segmentation process.



**Fig. 2.** LV segmentations (yellow) on a high-SNR cardiac MR image sequence. Ground truth shown in green.



**Fig. 3.** Comparison of the LV boundary estimates of our approach (yellow) with static segmentations using a shape prior (dark red) based on [8] on low-SNR data. Ground truth is shown in green.

## 6. CONCLUSION

We have presented an approach for segmentation of the evolving left ventricle. We have proposed a nonparametric, information-theoretic method to learn the time-varying LV dynamics from training data. Using a Bayesian framework, based on a single-cycle graph to capture periodicity of cardiac motion, we have developed a particle-based inference algorithm that allows time-recursive segmentation throughout the LV cycle using all available data. We have used finite-dimensional approximate representations of the LV boundary as state variables, rather than the infinite dimensional LV boundaries themselves, in order to improve the reliability and accuracy of the learning process. However, we have also incorporated curve evolution methods into this framework in order to generate boundaries not constrained to reside in the finite-dimensional domain in which dynamics have been learned and exploited. We have demonstrated the effectiveness of our approach (both qualitatively and quantitatively) on cardiac MR imagery. These results exhibit the improvements provided by our approach over static segmentation methods.

	Static Shape Prior	Proposed Approach
High SNR	0.8516	0.9292
Low SNR	0.8035	0.9069

**Table 1.** Dice coefficients based on test data consisting of 234 cardiac cycles of 20 frames each.

## 7. ACKNOWLEDGMENTS

The authors would like to thank Vivek Reddy and Godtfred Holmvang of the MR-CT Program at the Massachusetts General Hospital for providing the MR data.

## 8. REFERENCES

- [1] J. Duncan, A. Smeulders, F. Lee, and B. Zaret, "Measurement of end diastolic shape deformity using bending energy," in *Computers in Cardiology*, 1988, pp. 277–280.
- [2] J. C. McEachen II and J. S. Duncan, "Shape-based tracking of left ventricular wall motion," *IEEE Trans. Medical Imaging*, vol. 16, no. 3, pp. 270–283, 1997.
- [3] V. Chalana, D. T. Linker, D. R. Haynor, and Y. Kim, "A multiple active contour model for cardiac boundary detection on echocardiographic sequences," *IEEE Trans. Medical Imaging*, vol. 15, no. 3, pp. 290–298, 1996.
- [4] M-P. Jolly, N. Duta, and G. Funka-Lee, "Segmentation of the left ventricle in cardiac MR images," in *IEEE Int. Conf. on Computer Vision*, 2001, vol. 1, pp. 501–508.
- [5] J. Senegas, T. Netsch, C. A. Cocosco, G. Lund, and A. Stork, "Segmentation of medical images with a shape and motion model: A Bayesian perspective," in *Computer Vision Approaches to Medical Image Analysis (CVAMIA) and Mathematical Methods in Biomedical Image Analysis (MMBIA) Workshop*, 2004, pp. 157–168.
- [6] W. Sun, *Learning the Dynamics of Deformable Objects and Recursive Boundary Estimation Using Curve Evolution Techniques*, Ph.D. thesis, MIT, August 2005.
- [7] M. Leventon, *Statistical Models in Medical Image Analysis*, Ph.D. thesis, MIT, May 2000.
- [8] A. Tsai, A. Yezzi, W. Wells, C. Tempany, D. Tucker, A. Fan, W. E. Grimson, and A. Willsky, "A shape-based approach to the segmentation of medical imagery using level sets," *IEEE Trans. Medical Imaging*, vol. 22, no. 2, 2003.
- [9] D. Cremers, "Dynamical statistical shape priors for level set-based tracking," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 28, no. 8, pp. 1262–1273, Aug. 2006.
- [10] A. Ihler, "Maximally informative subspaces: Nonparametric estimation for dynamical systems," M.S. thesis, MIT, August 2000.
- [11] S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for on-line non-linear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Processing*, vol. 50, no. 2, pp. 174–188, 2002.
- [12] J. Pearl, *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufman, San Mateo, CA, 1988.
- [13] E. B. Sudderth, A. T. Ihler, W. T. Freeman, and A. S. Willsky, "Nonparametric belief propagation," in *IEEE Conf. on Computer Vision and Pattern Recognition*. IEEE, 2003.
- [14] L. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, pp. 297–302, 1945.