Graph-Driven Diffusion and Random Walk Schemes for Image Segmentation

Christos G. Bampis, *Student Member, IEEE*, Petros Maragos, *Fellow, IEEE*, and Alan C. Bovik, *Fellow, IEEE*

Abstract—We propose graph-driven approaches to image segmentation by developing diffusion processes defined on arbitrary graphs. We formulate a solution to the image segmentation problem modeled as the result of infectious wavefronts propagating on an image-driven graph, where pixels correspond to nodes of an arbitrary graph. By relating the popular susceptible-infected-recovered epidemic propagation model to the Random Walker algorithm, we develop the normalized random walker and a lazy random walker variant. The underlying iterative solutions of these methods are derived as the result of infections transmitted on this arbitrary graph. The main idea is to incorporate a degree-aware term into the original Random Walker algorithm in order to account for the node centrality of every neighboring node and to weigh the contribution of every neighbor to the underlying diffusion process. Our lazy random walk variant models the tendency of patients or nodes to resist changes in their infection status. We also show how previous work can be naturally extended to take advantage of this degree-aware term, which enables the design of other novel methods. Through an extensive experimental analysis, we demonstrate the reliability of our approach, its small computational burden and the dimensionality reduction capabilities of graph-driven approaches. Without applying any regular grid constraint, the proposed graph clustering scheme allows us to consider pixel-level, node-level approaches, and multidimensional input data by naturally integrating the importance of each node to the final clustering or segmentation solution. A software release containing implementations of this paper and supplementary material can be found at: http://cvsp.cs.ntua.gr/research/GraphClustering/.

Index Terms—Graph clustering, random walker, SIR epidemic propagation model, diffusion modeling, image segmentation.

I. INTRODUCTION

IMAGE segmentation is a process of dividing a digital image into smaller parts that provide meaningful insights about the objects and the structures in a scene. Segmentation provides useful information for many other higher level processes such as object detection or image retrieval. The traditional image segmentation literature is abundant in pixel-based methods, where every pixel is assigned to a particular label to create pixel segments. There are many categories of algorithms that are used in image segmentation such as edge-based [1], model-based [2], perception-based [3], region growing [4] methods etc. Some approaches do not assume user information (unsupervised), whereas others rely on it (interactive) by incorporating it as a hard or a soft constraint.

In this paper, we will focus on interactive image segmentation via graph-based clustering and segmentation. The main idea behind interactive segmentation is that the user marks a set of points (seeds) that he thinks belong to different objects. Then, this information is provided to a supervised algorithm that seeks to produce segmentation results that adhere well to these seeds. Some of the most influential interactive segmentation methods are based on graph cuts [5], normalized cuts [6] and the Random Walker (RW) algorithm [7], which is the focus of this work. In the RW setting, the user marks some pixels in the image. Then, assuming that a random walker starts from each unlabeled pixel, calculate the probability that these random walkers will first reach the already labeled pixels. At each pixel, the label with the greatest probability is picked as its final label yielding the final interactive segmentation result.

An important property of the RW algorithm is its connection to anisotropic diffusion [8], [9]. In particular, the RW algorithm shares the same energy functional as anisotropic diffusion and yields a steady state solution, whereas anisotropic diffusion provides a flow-based solution. By steady state solution we mean that the final solution takes a non-iterative closed form which usually is the solution of an appropriately defined linear system. This observation has been exploited to create alternatives to the RW approach. In order to handle natural images with complex textures, the Random Walker with restart (RWR) algorithm was introduced in [10]. In the RWR setting, the restarting probability c indicates that the random walker will return to the starting node (pixel) with a probability of c or walk out to an adjacent pixel with probability $1 - c$. Ham et al. [9] further extended the RWR to incorporate non-local information and structure.

In [11], a "lazy" random walk variant (LRW) was proposed and used for superpixel segmentation. A lazy random walker stays at the current node with probability $1 - \alpha$ and travels to some adjacent node with probability α. The LRW uses the normalized graph Laplacian and uses the notion of commute times instead of the first reach probability, like RW.

While the RW starts from the pixels to the seed points, the LRW algorithm computes the commute time from the seed points to other pixels. In [12], partially absorbing random walks (PARW) were used for video supervoxels. These approaches are typically called graph-based since they use a regular grid as the image representation domain.

One limiting property of the RW is that each segment has to be connected to a seed. In order to incorporate prior information into the RW approach and eliminate this constraint, Grady [13] proposed the use of an augmented graph where the additional nodes were connected to the original image's nodes. In a similar fashion, the recent work of Dong *et al.* [14], [15] proposed a unifying sub-Markov random walk approach (subRW). The subRW framework can be thought of as an extension of LRW, RWR and PARW where additional nodes are added in the base (original) image graph providing different kinds of image information. Other works have used spectral segmentation techniques based on the graph Laplacian [16]–[18]. For example, the Laplacian Coordinates (LC) method [16] considers an extended neighborhood at every pixel thereby improving the diffusive properties of this approach. An interactive segmentation approach was suggested in [19] that uses constrained Laplacian optimization along with an acceleration step that significantly reduces the computational time. A constrained version of RW that is suitable for multiple user inputs and that demonstrates improved results over RW was proposed in [20] while [21] used an error-tolerant graph cut energy function. In [22], the Power Watershed framework was developed to unify RW with shortest path optimization and graph cuts. In [23], a self-diffusion operator was introduced both for image segmentation and graph clustering and was related to random walks, but the authors did not focus on interactive image segmentation.

The common ground between all those approaches is that the base image graph is usually a regular image grid. While a regular grid can be an attractive option for devising different RW variants or for adding additional nodes (as in subRW), the regularity constraint on the image graph can be limiting for the following reasons. First, the developed algorithms rely on the simple image grid and do not consider other possible sources of information. Further, these approaches cannot be directly extended to other types of visual data such as point clouds or video sequences. This is not the case for arbitrary graphs which capture additional properties of the problem.

We believe that improved algorithms on image-driven graphs should not consider only simple 4- or 8-neighborhoods. Instead, we propose a set of RW approaches where pixels no longer have a one to one correspondence to nodes. It may occur that a single pixel is also a node, but most often a node represents a set of pixels in a small homogeneous region. In order to study interactions between the different nodes, we found it instructive and highly intuitive to relate our work to epidemic propagation models since the transmission of infections is also a kind of interaction between entities. Such models can be a source of inspiration when devising new RW algorithms, which are related to diffusion. Therefore, our first step is to relate a popular epidemic propagation model

called the Susceptible - Infected - Recovered (SIR) model to the RW by analyzing the transmission processes that occur between the nodes of an arbitrary graph.

We then take a step further by observing that the importance of a node with respect to the whole graph is a rich source of information. From a more abstract point of view, we can think of the node importance as the number of neighbors some node (or a human) has i.e. we can associate node importance to the notion of node centrality as with social graphs [24]. We incorporate this information into different diffusion schemes and interpret their steady state solutions as the final image (or graph) segmentation result. We introduce the Normalized Random Walker (NRW) and the Normalized Lazy Random Walker (NLRW) as those steady states. The term normalized refers to the fact that the underlying smoothness functional relates each node's degree with the degrees of its neighbors.

Previous methods such as subRW and RWR do not take into account this information. Further, LRW also uses the normalized graph Laplacian (like NRW) but the choice of the normalized Laplacian was "to be more consistent with the eigenvalues of adjacency matrices in spectral geometry and in stochastic process" [11]. Here, we discuss the effect of using the normalized Laplacian in great detail from both theoretical and practical perspectives. In addition, the proposed NRW and NLRW schemes decompose the Laplacian matrix (which is different from the one used in RW) into four sub-matrices and then convert those into a linear system. This idea originates from the original RW work and is conceptually different from the LRW that solves the inverse of the normalized Laplacian matrix to calculate the commute time.

In our previous work [25], we carried out preliminary studies on using NRW to improve the original RW algorithm. Going beyond our previous efforts, we further develop NRW and NLRW by studying both iterative and steady state solutions. We also demonstrate how to incorporate the node centrality term in other RW variants, such as subRW and RWR, which demonstrates that our approach is applicable in the design of other RW variants. Further, we carry out both pixel- and graph-based experiments and an extensive statistical analysis on many different datasets for the node-based case. The versatility of the proposed framework allows us to treat any 2D image as an arbitrary entity with irregular image patches. Therefore, it can be used not only for interactive image segmentation but also on other types of visual data such as video and point clouds.

The remainder of this paper is organized as follows. Section II provides an analysis of the SIR model. Then, in Section III, we briefly discuss the original RW algorithm and its main properties and drawbacks. A unification of the SIR model with the RW algorithm is presented in Section IV and Section V describes the NRW algorithm. In Section VI we analyze and extend the diffusion process related to the NRW and develop NLRW as its steady state solution. Section VII describes our full graph-based approach and Section VIII demonstrates the connections of NRW to subRW and RWR and ways to extend them. Section IX provides extensive results of our method compared to other seeded
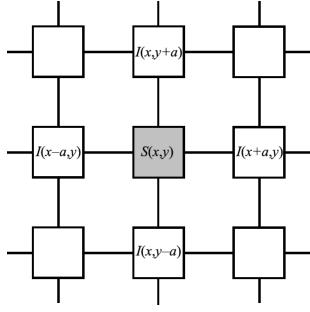
Fig. 1. SIR [27] regular grid around (x, y).

image segmentation methods. Finally, Section X demonstrates an application to visual point cloud and Section XI gives conclusions.

II. SUSCEPTIBLE INFECTED RECOVERED (SIR) MODEL

The SIR model [26], [27] is a well-studied epidemic propagation model in the field of mathematical epidemiology. The standard SIR (Kermack-McKendrick) model [26] considers a community of people belonging in 3 states: the susceptible (S), the infected (I) and the recovered (R). The model assumes that a recovered person can never be infected again and is described by the following three equations:

$$\frac{dS}{dt} = -kSI, \quad \frac{dI}{dt} = kSI - \frac{1}{\tau}I, \quad \frac{dR}{dt} = \frac{1}{\tau}I \qquad (1)$$

where k is the rate of infection transmission and τ is the recovery time, i.e. the time it takes for a person to recover after being infected. The functions $S(t), I(t), R(t)$ denote the number of individuals in each state respectively over time t. The model further assumes that the total number C of individuals does not change, i.e. $S(t) + I(t) + R(t) = C \; \forall t$. Postnikov and Sokolov [27] modified the SIR model by introducing spatial dependency on a continuum. Given a community of immobile individuals, the infection mechanism is assumed to be through contact with direct neighbors. As depicted in Fig. 1, all nodes form a regular grid with a constant spacing a between them. Then, $I(x, y, t)$ denotes the probability that the node at (x, y) in the 2D regular grid will be infected at time t ($S(x, y, t)$ and $R(x, y, t)$ are similarly defined). The change over time $\Delta I(x, y, t)$ of the probability of being infected at (x, y) is modeled as:

$$\Delta I(x, y, t) = \frac{k}{4} S(x, y, t) \big[I(x + a, y, t) + I(x - a, y, t)$$
$$+ I(x, y + a, t) + I(x, y - a, t) \big] \Delta t \qquad (2)$$

Therefore, the contact mechanism assumes an average infection state over the 4-node neighborhood and an infection flow proportionate to $\frac{k}{4}$ over all orientations provided that the node is susceptible to the infection. In addition, it is proportionate to the change in time.

III. RANDOM WALKER ALGORITHM

The diffusive nature of the SIR model in (2) may be used for image segmentation via random walk theory. In [7], Grady proposed the Random Walker (RW) algorithm

for image segmentation, where the user inputs a set of seeds, each corresponding to one of M possible segmentation labels. This set of marked pixels is then used as the initial state towards extracting the desired object boundaries. The final segmentation is interpreted as the most probable label at each node, i.e. which type of seed is the most probable final destination for a random walker who starts his trip from every unmarked node. However, direct solution of the random walker's probabilities with respect to all unmarked nodes is computationally intractable [7]. Instead, one may solve the relevant Dirichlet problem (also known as the combinatorial Dirichlet integral for graph applications) [28], [29] and then apply the RW algorithm. We now briefly describe the RW algorithm. Let $\mathbf{L}$ be the unnormalized graph Laplacian [7]:

$$\mathbf{L} = \mathbf{D} - \mathbf{W} = [l_{ij}] = \begin{cases} d_i, & \text{if } i = j \\ -w_{ij}, & \text{if } j \sim i \\ 0, & \text{else} \end{cases} \qquad (3)$$

where N is the number of pixels in the image, $\mathbf{W} = [w_{ij}]$ is the weight matrix, $\mathbf{D} = \text{diag}(d_1, \ldots, d_N)$ is the degree matrix, $d_i = \sum_{j \sim i} w_{ij}$ is the degree of node i and $\sim$ denotes that pixels or nodes i and j are adjacent. To construct the weights, a gaussian kernel was proposed using pixel luminances (or colors) $\mathbf{g}_i$ at each pixel i, i.e. $w_{ij} = \exp(-\|\mathbf{g}_i - \mathbf{g}_j\|^2 / \sigma^2)$ where σ is a scale-related parameter of the algorithm. Then we minimize

$$J(\mathbf{x}) = \frac{1}{2} \sum_{i \sim j} w_{ij} \big(x_i - x_j \big)^2 = \frac{1}{2} \mathbf{x}^\top \mathbf{L} \mathbf{x} \qquad (4)$$

given the constraints (seed locations) encoded in $\mathbf{x} = [x_i]$ which are the random walker's probabilities for one of the M different seed types (e.g. for $M = 2$ we have a set of seeds for the background and a set of seeds for the foreground of the image). Then, at every unmarked node, we choose the label that corresponds to the maximum probability i.e. we find the largest of the M different x_i values (one for each seed type). The RW algorithm outputs the labels assigned to every pixel. Meanwhile, the random walker's probabilities have three important properties:

1) they sum to one across all possible labels
2) they are between zero and one
3) they are equal to a linear combination of all the values in the neighborhood i.e. $x_i = \frac{1}{d_i} \sum_{i \sim j} w_{ij} x_j$

Properties 1 and 2 reflect the underlying probabilistic interpretation of the random walker and properties 2 and 3 are combinatorial analogues of properties of continuous harmonic functions [7]. In fact, it is well known that the RW solution can be regarded as the steady state solution of the heat equation: $\frac{du}{dt} = \nabla^2 u$. This equation models the heat flowing from "hot" to "cold" areas in a graph [7], [30]. Similarly, it can be viewed as a label propagation approach for graph-based learning models [31]. Note that we are minimizing the functional $J(\mathbf{x})$ for each of the M seed types thus it is required to solve a set of M systems of linear equations [7]. However, if one uses property 1 of the generated probabilities, then $M - 1$ systems of linear equations need to be solved.
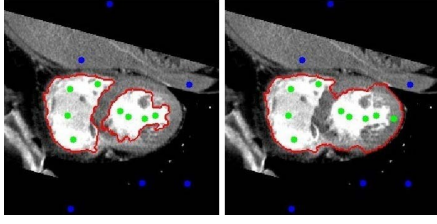
Fig. 2.   RW is sensitive to the number and location of the seeds.

As in most seeded segmentation methods, the RW algorithm is sensitive to the initial seed location, number and quality, e.g. the results may be adversely affected if the human factor introduces incorrect initial seeds due to lack of expertise or concentration (see Fig. 2). In [20], the authors proposed an improved way of picking the seeds that delivers improved results compared to the RW algorithm. Parameter tuning is also an important aspect, since prior knowledge of σ is usually unavailable and can vary from image to image. Searching for an optimal σ iteratively or by an explicit formula as in [32] digests precious computational time. In the experimental section we study the robustness of our proposed NRW method against the RW algorithm and other methods, while varying the parameter σ, for different locations and number of seeds.

IV. Unification of the SIR With the Random Walker Algorithm

Next, we derive the RW solution as the steady state solution of the SIR diffusion. Our ultimate goal is to model graph clustering as the steady state of different infections whose status increases at already infected nodes and propagates to uninfected ones through the contact spread mechanism of the SIR. Consider every image pixel to be equally susceptible to infections/labels diffused by the SIR model, i.e. $S(x, y, t) = 1 \ \forall x, y, t$. Also assume that there are no individuals in the recovery state, i.e. let $\tau \to \infty$. We also normalize the time steps so that $\Delta t = 1$. Then, the SIR model becomes:

$$\Delta I(x, y, t)$$
$$= \frac{k}{4}\big[I(x + a, y, t) + I(x - a, y, t) + I(x, y + a, t)$$
$$+ I(x, y - a, t)\big] = I(x, y, t + 1) - I(x, y, t)$$

Then, given an initial infection state, the infection probability at (x, y, t) is updated, i.e.

$$I(x, y, t + 1) = I(x, y, t) + \frac{k}{4}\big[I(x + a, y, t)$$
$$+ I(x - a, y, t) + I(x, y + a, t) + I(x, y - a, t)\big] \quad (5)$$

Define the initial state as follows: suppose there exists a set $P = \{s_k\}$ of breakout points located at $s_k = (x_k, y_k)$ related to some infection. We rewrite $I(x_i, y_i, t)$ as $I_{i,t}$ to denote the infection probability at time t of pixel/node i located at (x_i, y_i). Similar to other seeded segmentation methods, consider these points to have reference values for their infection status, i.e. $I_{i,t} = 1 \ \forall t$ for some infection or label and zero everywhere else. Accordingly, all other points are initially at

a healthy state for all infections, i.e. $I_{i,0} = 0 \ \forall i \in N_o \setminus P$, where $N_o = [1, \ldots, N]$ and N is the number of nodes.

Next, reformulate (5) to express the increase in infection in terms of the contact mechanism. A real life analog could be that friendships among members of a community share more experiences and are more prone to be infected by someone they spend more time with. To model node similarity we use weights w_{ij}, where i, j denote two pixels/nodes in the image/graph. Constructing the weight matrix $\mathbf{W}$ is important for the final segmentation: higher w_{ij} means higher similarity between nodes i and j thus a stronger connection between them. We use a gaussian kernel with parameters σ_g, σ_h:

$$w_{ij} = \begin{cases} \exp\left[-\dfrac{\|\mathbf{g}_i - \mathbf{g}_j\|_2^2}{\sigma_g^2} - \dfrac{\|\mathbf{h}_i - \mathbf{h}_j\|_2^2}{\sigma_h^2}\right], & \text{if } j \sim i \\ 0, & \text{else} \end{cases} \quad (6)$$

where $\mathbf{g}_i$ is the image feature vector (e.g. a 3D RGB vector) associated with pixel/node i, and $\mathbf{h}_i$ is the location of pixel/node i. Color features are often used in RW variants and filterbank features have been used in [33] for texture segmentation. Consider replacing k with w_{ij}, denote by d_i the degree of node i and reformulate (5) to obtain the steady state ($t \to \infty$) at some node i:

$$\Delta I_{i,t} = I(x_i, y_i, t + 1) - I(x_i, y_i, t) = \sum_{j \sim i} \frac{w_{ij}}{d_i} I_{j,t}$$

$$\overset{\substack{\Delta I_{i,t} \to 0 \\ t \to \infty}}{\Longrightarrow} \sum_{j \sim i} \frac{w_{ij}}{d_i} I_{j,t} = 0 \quad (7)$$

where d_i is the degree of node i. For $w_{ij} > 0$ and $I_{j,t} \geq 0$ (7) is not applicable, i.e. it implies $I_{j,t} = 0 \ \forall j, t$. Therefore, we reformulate the SIR by assuming the following local mean field model [27]:

$$I_{i,t+1} = I_{i,t} + \sum_{j \sim i} \frac{w_{ij}}{d_i}(I_{j,t} - I_{i,t}), \quad (8)$$

where by node i becomes more infected as time passes. Note that the amount of neighbor influence is proportional to the difference between each pair's infection status. When node i and j are similarly affected, then j ceases to affect i. The contact mechanism is symmetric: i has the same effect on j, hence the weight matrix $\mathbf{W}$ is symmetric. This formulation can also be interpreted through the use of the derivative $\partial_j I_{i,t} = \sqrt{w_{ij}}(I_{j,t} - I_{i,t})$ along edge e_{ij} [32] (see Appendix). Note that (8) holds for all propagating infections evolving as $\Delta I_{i,t} = I_{i,t+1} - I_{i,t}$. When a steady state is reached, then:

$$\Delta I_{i,t_\infty} = \sum_{j \sim i} w_{ij}(I_{j,t_\infty} - I_{i,t_\infty}) = 0$$

$$\Rightarrow I_{i,t_\infty} = \sum_{j \sim i} \frac{w_{ij}}{d_i} I_{j,t_\infty} \quad (9)$$

where t_∞ means that $t \to \infty$. Therefore, the steady state corresponds to cessation of evolution (updating) of the infection status at all nodes and for all infections. Also, I_{i,t_∞} satisfies the previous (harmonic) properties 2 and 3 [7], [30]:
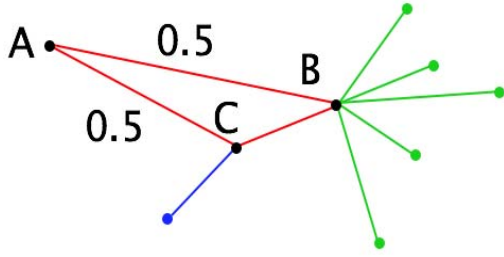
Fig. 3. A has 2 connections, B has 7 and C has 3, colors denote independent neighborhoods.

1) $0 \leq I_{i,t_\infty} \leq 1$ at all nodes and infections
2) I_{i,t_∞} is equal to the weighted average of its neighboring nodes (the mean-value theorem)

Our reformulation of the SIR model may now be interpreted in terms of the RW algorithm: if one considers labels as different infections and the probabilities x_i as I_{i,t_∞} in (9), then the steady state of the extended SIR model is identical to the solution provided by the RW algorithm. Note that the weighted average property of the reformulated SIR is equivalent to the third property of the RW by simple inspection. Motivated by this natural connection, we further develop the Random Walker algorithm in new and different ways by using arbitrary graph structures and node networks. In our SIR reformulation, we did not take into account the dynamics of the susceptible and recovered compartments. For image segmentation tasks, it is more useful to consider the interactions between pixels/nodes thus we designed the SIR such that each node is infected by all types of infections and the final segmentation result (which is the desired output for RW) is expressed as finding the infection that dominated over each node in the end i.e. when the steady state of the SIR diffusion process is reached.

V. NORMALIZED RANDOM WALKER (NRW)

By analyzing (8), it may be observed that the nodal degree is not accounted for when computing the edge derivative. However, this would not be natural in a realistic disease propagation: the number of neighbors each person has influences the local infection profile. Suppose that A has 2 friends, B and C (see Fig. 3). Also, suppose that both of them are "equal" friends of A, but B has more friends than C. Then, B will more heavily influence A's infection profile. In other words, the probability of a node's infection is related to its degree and that of its neighbors. We thus incorporate degree terms to satisfy this observation. Denote the normalized graph Laplacian [34] by $\tilde{\mathbf{L}} = \mathbf{D}^{-\frac{1}{2}}(\mathbf{D} - \mathbf{W})\mathbf{D}^{-\frac{1}{2}} = [\tilde{l}_{ij}]$, where

$$\tilde{l}_{ij} = \begin{cases} 1, & \text{if } i = j \\ -\dfrac{w_{ij}}{\sqrt{d_i d_j}}, & \text{if } j \sim i \\ 0, & \text{else} \end{cases} \quad (10)$$

and $\mathbf{W}, \mathbf{D}$ are the weight and degree matrix respectively. Also, define the normalized derivative [32] over edge e_{ij} as $\partial_j I_{i,t} = \dfrac{\sqrt{w_{ij}}}{\sqrt{d_i}}\left(\dfrac{I_{j,t}}{\sqrt{d_j}} - \dfrac{I_{i,t}}{\sqrt{d_i}}\right)$. Our new iterative scheme is then written

as (see Appendix):

$$I_{i,t+1} = I_{i,t} + \sum_{j \sim i} \frac{w_{ij}}{\sqrt{d_i}}\left(\frac{I_{j,t}}{\sqrt{d_j}} - \frac{I_{i,t}}{\sqrt{d_i}}\right) \quad (11)$$

Proposition 1: The iterative scheme (11) can be re-written as:

$$\mathbf{\Delta I}_t = -\tilde{\mathbf{L}}\mathbf{I}_t \quad (12)$$

where $\mathbf{\Delta I}_t$ and $\mathbf{I}_t$ are $N \times 1$ vectors and N is the number of graph nodes or image pixels:

$$\mathbf{\Delta I}_t = (\Delta I_{1,t} \ldots \Delta I_{N,t})^\top, \quad \mathbf{I}_t = (I_{1,t} \ldots I_{N,t})^\top \quad (13)$$

Also, the steady state solution of (11) minimizes:

$$J_n(\mathbf{x}) = \frac{1}{2}\sum_{j\sim i} w_{ij}\left(\frac{x_i}{\sqrt{d_i}} - \frac{x_j}{\sqrt{d_j}}\right)^2 \quad (14)$$

which is a normalized version of $J(\mathbf{x})$ in (4). As shown in the Appendix (proof of Proposition 1), minimizing $J_n(\mathbf{x})$ requires solving an appropriately defined linear system of equations.

Proposition 2: The steady state at node i is:

$$I_{i,t_\infty} = \frac{1}{\sqrt{d_i}} \sum_{j \sim i} \frac{w_{ij}}{\sqrt{d_j}} I_{j,t_\infty}. \quad (15)$$

Proposition 3: I_{i,t_∞} violates the maximum principle of the RW, hence M systems of linear equations must be solved, one for each label (or infection type). We derive the proofs of Proposition 2 and 3 in the Appendix.

VI. NORMALIZED LAZY RANDOM WALKER (NLRW)

Neither RW nor NRW consider the resistance of nodes to changes. For example, an infected patient may attempt to reduce his chances of getting sicker by changing his daily habits, by using medications etc. We seek to incorporate this behavior by extending (11). Similar to [11], we introduce a free parameter $\alpha \in [0, 1]$ which expresses the probability of the random walker making a self-loop. We then propose the following iterative scheme:

$$I_{i,t+1} = I_{i,t} - \frac{\alpha d_i}{1 - \alpha + \alpha d_i} I_{i,t} + \alpha \sum_{j \sim i} \frac{w_{ij}}{\sqrt{d_i}} \frac{I_{j,t}}{\sqrt{d_j}} \quad (16)$$

Let $\mathbf{W}' = [w'_{ij}]$ denote the weight matrix of the lazy random walk:

$$w'_{ij} = \begin{cases} 1 - \alpha, & \text{if } i = j \\ \alpha w_{ij}, & \text{if } j \sim i \\ 0, & \text{else.} \end{cases} \quad (17)$$

Then $\mathbf{D}' = \text{diag}(d'_1, \ldots, d'_n)$, where $d'_i = 1 - \alpha + \alpha d_i$. Similar to Proposition 2, the steady state of (16) corresponds to

$$I_{i,t_\infty} = \frac{1 - \alpha + \alpha d_i}{d_i} \frac{1}{\sqrt{d_i}} \sum_{j \sim i} \frac{w_{ij}}{\sqrt{d_j}} I_{j,t_\infty} \quad (18)$$

$$= \frac{d'_i}{d_i} \frac{1}{\sqrt{d_i}} \sum_{j \sim i} \frac{w_{ij}}{\sqrt{d_j}} I_{j,t_\infty} \quad (19)$$
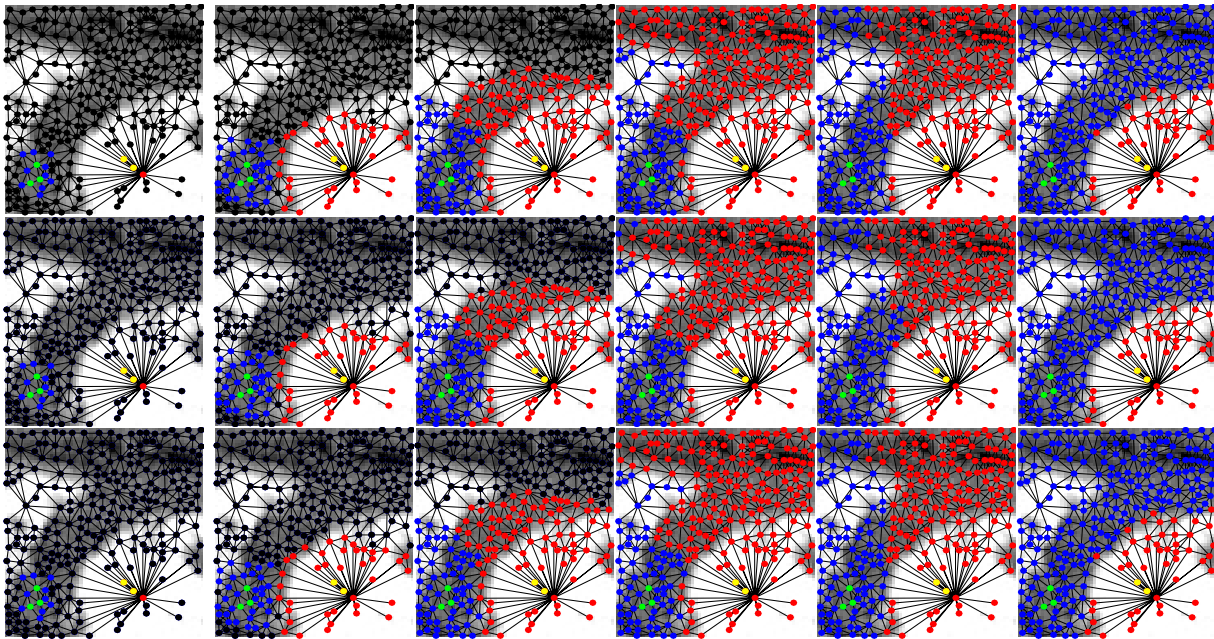
Fig. 4.   Columns 1 - 5: Transmission of $M = 2$ infections, timesteps (left to right): 1, 2, 5, 50 and 5000. Column 6: Steady states, rows 1 - 3: RW, NRW and NLRW. Blue nodes: Infection 1. Red nodes: Infection 2. Black nodes are not yet infected, while seeds denoted by green and yellow denote infection 1 and infection 2 respectively.

Note that (19) degenerates to the NRW solution when $\alpha = 1$. When $\alpha = 0$ the node remains in the previous state since the probability of a self-loop is one. We call the steady state solution of (16) NLRW to avoid confusion with the LRW method [11]. As with RW and NRW, the NLRW algorithm updates the infection probability function $I_{i,t}$ at every unseeded node and for every infection. All these update rules can be written in an abstract matrix form as $\Delta \mathbf{I}_t = -\mathbf{L}\mathbf{I}_t$, each with a different Laplacian. In the case of NLRW, the graph Laplacian $\mathbf{L}$ is denoted by $\mathbf{L}' = [l'_{ij}]$ and is given by:

$$
l'_{ij} = \begin{cases} \dfrac{\alpha d_i}{1 - \alpha + \alpha d_i}, & \text{if } i = j \\ -\dfrac{\alpha w_{ij}}{\sqrt{d_i}\sqrt{d_j}}, & \text{if } j \sim i \\ 0, & \text{else.} \end{cases} \tag{20}
$$

Note that using $\mathbf{L}' = \mathbf{D}' - \mathbf{W}'$ for NLRW is not useful, since in that case $\mathbf{L}' = \alpha \mathbf{L}$ which produces the same solution as (8). We verified this observation empirically. Our NLRW scheme is different from LRW since it does not use the notion of commute times [11], it is not constrained to a pixel-based framework and it is strongly related to the NRW scheme. Fig. 4 shows the evolution of the underlying diffusion schemes for RW, NRW and NLRW.

VII. Graph-Based Normalized Random Walker

Next, we present a NRW model for arbitrary graphs. Let $G = (V, E)$ denote an image-driven graph consisting of a set of vertices (nodes) $v \in V$ and a set of edges $e \in E \subseteq V \times V$. In addition, an edge e_{ij} denotes an edge spanning two vertices v_i and v_j. In order to represent node similarity, we define non-negative weights w_{ij} between nodes i and j and the
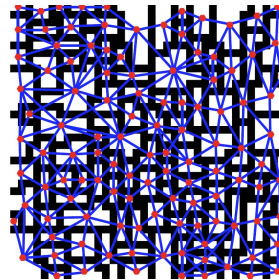

Fig. 5.   Region Adjacency Graph (RAG) based on watershed.

degree d_i as before. Also assume that G is connected and undirected ($w_{ij} = w_{ji}$). Instead of applying our algorithm on a pixel level (regular grid), we extend it to a Region Adjacency Graph (RAG) [35]. Then we can use any over-segmentation technique (such as a watershed transformation [36] or SLIC [37]) to obtain a set of n regions $R_1, R_2, ..., R_n$ represented by a node i located at the spatial mean $\mathbf{h}_i$ of R_i and whose feature vector is $\mathbf{g}_i$, i.e. the mean feature vector over all pixels in R_i (see Fig. 5). Consequently, the problem dimensionality can be greatly reduced. We prefer a node-level approach, since it better approximates real-life situations, where the number of friends or neighbors of each person is free of spatial regularity. From an image segmentation perspective, the degree of each node will more accurately depict the relationship between adjacent image regions, thus capturing local image structures more efficiently.

A drawback of RAG is that it is necessary to interpret the initial seeds in terms of their corresponding nodes. A good approximation can be obtained by assigning each seed's label to the node closest to the seed's pixel, by assuming that the user will not input any seeds very close to an object's
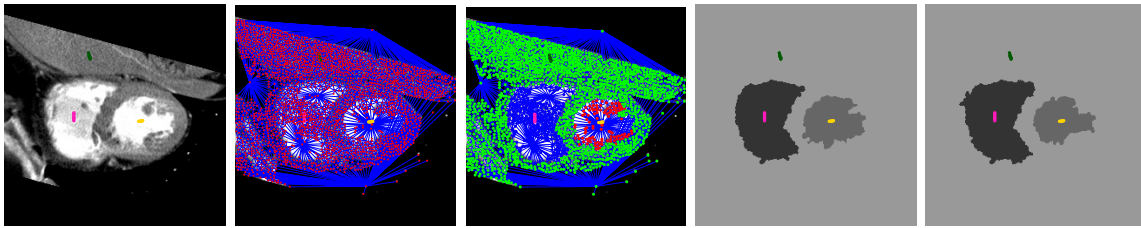
Fig. 6. From left to right: CT image + seeds, image-driven graph, graph clustering in 3 classes using the RW algorithm, corresponding pixel-level RW and pixel-level NRW using the graph clustering solution.

boundary, i.e., within a few pixels, thus training the human user is not necessary. A more accurate (but slower) approach would be to assign the seed's label to the region or node the pixel belongs to. Also, to transform the nodal solution into a spatial one, assign the label of node i to all pixels in region R_i. Finally, boundary pixels are handled by assigning each to the most common label present around its eight neighbors. Fig. 6 shows the previously described steps.

Using an arbitrary number of neighbors can be advantageous for NRW. Intuitively, NRW exploits the node degree information which is less informative when the number of neighbors is restricted to 4 or 8. Given the fundamental property that real-world images are locally correlated, this implies a very strong advantage for NRW in its node-based approach. Also, the use of the normalized graph Laplacian aligns with the theoretical merits described in [38].

VIII. EXTENSIONS AND CONNECTIONS WITH PREVIOUS METHODS

As discussed earlier, NRW minimizes $J_n(\mathbf{x})$ which corresponds to re-weighting of a smoothness term which ensures that if nodes i and j are similar (high w_{ij}), then the corresponding probabilities x_i and x_j are as close as possible. A trivial solution is to set all $x_i = 0$. However, in the case of NRW, minimizing $J_n(\mathbf{x})$ is also constrained by the seed information. This observation leads us to connect and extend NRW to other RW variants such as the subRW method [14]. We only give a high-level overview of the subRW in order to connect the ideas of NRW with many other RW variants (for this section only we adopt the exact same notation of [14] to avoid confusion). Consider the subRW functional:

$$\bar{O}^{l_k}$$

$$= \frac{1}{2} \sum_{i \sim j} w_{ij} \left(\bar{r}_{im}^{l_k} - \bar{r}_{jm}^{l_k} \right)^2 + \frac{1}{2} \sum_{i=1}^{N} \frac{(d_i + \lambda g_i) c_i}{1 - c_i} \left(\bar{r}_{im}^{l_k} - b_{im}^{l_k} \right)^2$$

$$+ \frac{1}{2} \sum_{i=1}^{N} \lambda u_i^k \left(\bar{r}_{im}^{l_k} - 1 \right)^2 + \frac{1}{2} \sum_{t=1, t \neq k}^{K} \sum_{i=1}^{N} \lambda u_i^t \bar{r}_{im}^{l_k^2}$$

where $\bar{r}_{im}$ is the labeling probability of node i for the $m-$ seeded pixel, c_i is the restarting probability of node i, k is the label index for label l_k, λ is a constant controlling the effect of a GMM label prior and $b_{im}^{l_k}$ is related to the seed information. The first term corresponds to the smoothness term between the labeling probabilities, the second controls the effect of the seeded information and the last two correspond to

the GMM prior encoded in u_i^k. More importantly, the subRW is related to many other RW variants. For example, setting $\lambda = 0$ (no GMM prior) and $c_i = c \; \forall \; i$ yields the RWR.

Clearly, subRW uses a smoothness term that is not incorporating the degree-aware term of NRW. It is possible to re-express subRW so that the relative degree between the nodes is taken into account. Indeed, the smoothness term of NRW may be quite useful when designing other RW variants, since it can be naturally integrated into the smoothness term of any optimization function. In the Appendix, we show two examples of this integration. Integrating the degree information is not bound to any assumptions and can be successfully deployed in other methods as we demonstrate next.

IX. EXPERIMENTAL RESULTS

We evaluated both pixel-based and node-based approaches to demonstrate the power of the NRW degree-aware term.

A. Pixel-Based Experiments

For the pixel-based experiments we considered the following methods: RW [7], LC [16], LRW [11], RWR [10], subRW [14] and our proposed new models. For all methods we used publicly available implementations. The most popular dataset for validating interactive segmentation results is the GrabCut Microsoft Research (MSRC) dataset [39], which includes 50 images, their both ground truth segmentations, and their tri-maps. We used four commonly used comparison measures [16]: the Rand Index (RI), the Global Consistency Error (GCE), the Variation of Information (VoI) and the error rate. The RI measures how closely the resulting segmentation and the ground truth agree by counting the number of corresponding pixel pairs having the same labels. Therefore, higher results correspond to better segmentation results. The GCE quantifies the level of refinement between two segmentations while the VoI captures the distance between the segmentations in terms of their relative entropies. Lower values of both the GCE and the VoI indicate better segmentation results. Finally, the error rate is the percentage of misclassified pixels, thus lower values are better.

For a fair comparison between the different approaches, we conducted an extensive parameter search for each method. For RW, LC and NRW it is necessary to determine σ_g only. For LRW and NLRW both σ_g and α must be found, while for RWR, subRW and NsubRW we must find σ_g and c. For the latter two, we set λ to the default value suggested in [14]. Next, we fixed all parameters to yield the best RI value,

TABLE I

GRABCUT MSRC DATASET [39]: RI, GCE, VoI AND ERROR %; PIXEL-BASED RESULTS AFTER DETERMINING THE BEST SET OF PARAMETERS FOR
EACH METHOD. FOR EACH METRIC, THE FIRST ENTRY CORRESPONDS TO THE AVERAGE VALUE AND THE SECOND TO THE MEDIAN

Method	RI($\uparrow$)		GCE($\downarrow$)		VoI($\downarrow$)		error % ($\downarrow$)		time (sec.)	best σ_g	best α	best c
RW [7]	0.9703	0.9749	0.0275	0.0235	0.1911	0.1712	5.1671	5.0136	0.4193	70	-	-
LC [16]	0.9765	0.9799	0.0219	0.0196	0.1589	0.1484	4.1684	3.9260	3.1689	22	-	-
LRW [11]	0.9774	0.9824	0.0216	0.0173	0.1575	0.1373	3.8900	3.3468	2.4602	20	0.95	-
RWR [10]	0.9812	0.9881	0.0177	0.0118	0.1314	0.0993	3.1127	2.3324	1.7555	10	-	0.5
subRW [14]	0.9812	0.9881	0.0177	0.0118	0.1314	0.0992	3.1125	2.3314	50.1021	10	-	0.5
NRW	0.9750	0.9802	0.0234	0.0193	0.1684	0.1498	4.3463	3.8572	0.6049	100	-	-
NLRW	0.9749	0.9801	0.0235	0.0194	0.1691	0.1502	4.3676	3.8347	0.6213	100	0.9999	-
NsubRW	0.9813	0.9882	0.0176	0.0117	0.1309	0.0987	3.0978	2.2755	56.8666	10	-	0.5

TABLE II

LEFT: PIXEL-BASED RESULTS ON GRABCUT MSRC DATASET [39]. RIGHT: DATASET CHARACTERISTICS FOR GRAPH-BASED EXPERIMENTS.
(a) RI, GCE, VoI AND ERROR % ON GRABCUT MSRC; PIXEL-BASED RESULTS ON SEED SETS S_1 AND S_2 AFTER DETERMINING
THE BEST SET OF PARAMETERS FOR EACH METHOD USING THE TRIMAPS. (b) DATASET CHARACTERISTICS

(a)

Set	S_1				S_2			
Method	RI($\uparrow$)	GCE($\downarrow$)	VoI($\downarrow$)	error % ($\downarrow$)	RI($\uparrow$)	GCE($\downarrow$)	VoI($\downarrow$)	error % ($\downarrow$)
RW [7]	0.8226	0.1147	0.6793	5.5812	0.8987	0.0809	0.4668	2.9118
LC [16]	0.8266	0.1241	0.7027	5.3715	0.8727	0.0982	0.5664	3.7538
LRW [11]	0.7849	0.1745	0.9308	6.4207	0.8006	0.1568	0.8520	6.1197
RWR [10]	0.7785	0.1803	0.9539	6.6478	0.7914	0.1635	0.8843	6.4394
subRW [14]	0.7955	0.1622	0.8789	6.0712	0.8095	0.1489	0.8158	5.8058
NRW	0.8436	0.1081	0.6273	4.7679	0.9129	0.0723	0.4209	2.4782
NLRW	0.8420	0.1091	0.6323	4.8030	0.9125	0.0726	0.4224	2.4923
NsubRW	0.7956	0.1624	0.8795	6.0651	0.8095	0.1490	0.8161	5.8067

(b)

Dataset	# images	# used	# labels	# seeds
W_1	100	100	1	15
W_2	100	97	2	15
S	100	100	≥ 1	20...50
G_1, G_2	50	50	≥ 1	20...50
V	2913	2798	≥ 1	15
E	1000	999	≥ 1	15...40

then computed the average of all four evaluation metrics over all images. The search space for each parameter was: $\sigma_g \in [10, 140]$, $\alpha \in [0.95, 0.9999]$, $c \in [0.0001, 0.5]$. All results are tabulated in Table I. Clearly, RW was the worst performing algorithm, followed by NLRW and NRW which were slightly worse than LC; LRW is slightly better than LC but worse than RWR, subRW and NsubRW. Among the rest of the methods it is hard to reach any conclusions since the average differences were similar. RWR, subRW and NsubRW yielded the best numerical performance but these differences were not statistically significant. Alternatively, we could have compared RWR, subRW and NsubRW by computing the median value of the error rate over all pairs [σ_g, α], giving 4.9527, 4.8782 and 4.5308 for RWR, subRW and NsubRW respectively, a slight advantage for NsubRW.

To further examine performance in another light, we used two other sets of seeds proposed in [40] which we call seed set 1 (S_1) and seed set 2 (S_2). These seed points were more sparse than the original seed points in the GrabCut dataset. The second set (S_2) contained more labeled points than the first set, thus the results across all methods would likely be better for S_2. Next, fixing the parameters computed using the original seed points from the GrabCut dataset, we tested on S_1 and S_2. The quantitative results are tabulated in Table II(a), which shows that RWR, subRW and NsubRW delivered much lower performance than before, while NRW outperformed them. It is likely that the parameters chosen for the trimaps were not optimal for the seeds in S_1 and S_2. Further, subRW and NsubRW gave nearly identical performance. This may be due to the fact that for more complex RW variants (like subRW)

the degree-aware term may not always add more information. Meanwhile, NRW is a reliable alternative since the degree-aware term is not affected by any parameter choice. For S_1 and S_2 NRW performed the best compared to all other methods.

In Table I, we also include the average compute time over 50 images on the GrabCut dataset. Clearly, subRW and NsubRW are costly, due to the GMM prior computation, while NsubRW is even more so since it computes the normalized graph Laplacian. On the other hand, RW NRW and NLRW are very fast owing to their simplicity. The slight increase for both NRW and NLRW compared to RW is again due to the computation of $\mathbf{D}^{-\frac{1}{2}}$. However, this slight increase in compute time is outweighed by the improved segmentation quality indicated by the segmentation quality metrics. The LC method and LRW are more expensive since they involve more complex computations.

Finally, we visually compared the different methods using either the trimaps from the GrabCut dataset (row 1) or some input scribbles (row 2 and 3) as shown in Fig. 7. Regarding row 1, LRW performed the best, while NRW was the second best approach. When comparing subRW and NsubRW, NsubRW was much better, though still inferior to LRW. This is due to the fact that for every image the best value for subRW (and NsubRW) may vary and for this example, the default parameters of subRW may have been suboptimal. This visual example also suggests that while subRW and NsubRW may perform equally well when only the best set of parameters is considered, NsubRW is more robust against parameter choice. This is a highly desirable property for any

Fig. 7. Comparison between pixel-based methods. Row 1: $\sigma = 60$, $\alpha = 0.99$, $c_{\text{RWR}} = c_{\text{subRW}} = 10^{-4}$, $c_{\text{NsubRW}} = 10^{-3}$; Rows 2 and 3: $\sigma = 90$, $\alpha = 0.99$, $c_{\text{RWR}} = c_{\text{subRW}} = 10^{-4}$. For each figure, the error rate (e) and the RI are reported for each method.

segmentation method. Regarding rows 2 and 3, NRW clearly performed better than all other methods both in terms of RI and the error %.

B. Arbitrary Graph-Based Experiments

In the pixel-based experiments, the performance of NRW (and NLRW) was reliable across all seed points and comparable to LC. We verified that node centrality information is not always a strong source of information for pixel-based methods, when the number of neighbors is 4 or 8. However, our ultimate goal is to deploy NRW in a graph-based scheme where the degree-aware term will likely contribute more.

In the node-based experiments we compared RW, LC, NRW, LRW, NLRW and RWR, using the publicly available implementations extended to be applicable to arbitrary graph schemes. We did not include subRW since it is pixel-based. We also tried the NRWR scheme mentioned in the Appendix, which is an extension of RWR and is also related to subRW. In all the experiments, NRWR improved upon RWR. This also suggests that integrating degree-aware information can improve a particular method. However, we observed that the performance of NRWR was upper bounded by LRW i.e. fixing $1 - c = \alpha$ yielded the best results. Therefore, we do not report results on NRWR.

We used six different datasets and their ground truths (see Table II(b) for a description). The datasets used were: Weizmann [41] one object (W_1) and two objects (W_2), Semantic100 [42] (S), GrabCut MSRC [39] (G_1 for the seeds in S_1 and G_2 for S_2), PASCAL VOC12 [43] (V) and ECSSD [44] (E). On the Weizmann dataset with two objects and the PASCAL VOC12 dataset, we removed a small percentage of the images (containing very small objects) since the number of nodes inside them was too small. This step did not affect the overall performance significantly for any method. Unfortunately, most of these datasets do not contain pre-specified seed points/pixels which would allow direct and fair comparisons, hence we took a two-step approach. First, the

seed set S_1 in the GrabCut dataset was used to find the best set of parameters for each method. Then, given the ground truth labels for every other dataset, we drew random seed pixels which are included within each label. We set a constant number of random seeds without replacement so that they are unique within each region. These pixel seeds were then assigned to their nearest nodes, which now become seed nodes. To cover each labeled region accurately, we used a different number of seeds and many iterations in each experiment.
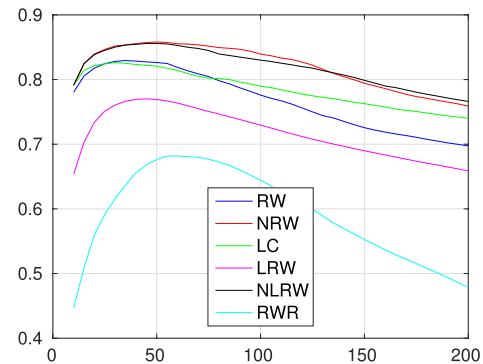
Next follows the node-based evaluation steps. First, the RAG was constructed for each test image and the weight matrix **W** computed using Lab colorspace features. These color features yielded improved performance compared to RGB values or grayscale values [25]. Then, each method was applied using the best of parameters from S_1 to produce the final labels for each node belonging in the RAG. Next, an appropriate metric was used to compare the labels derived by each method and the available ground truth masks. Given the RAG mapping, each pixel corresponds to some node depending on which region the pixel belongs to. This allows us to propagate pixel labels by assigning them to corresponding nodes. Then, the reference node labels are used to compare them with each node-level solution using the RI, GCE and VoI metrics.

In our first experiment, we studied the performance of each method across all datasets. Table III(a) tabulates the average RI results for all methods and datasets using 50 iterations. This averaging step was done over iterations within a single image and then over all images in each dataset. RWR gave the worst performance while that of LC and RW was similar. LRW was less competitive than in the pixel-based experiments, performing worse than RW. Clearly, the degree-aware term incorporated in NRW allowed the best result across all datasets. NLRW produced results very similar to NRW. This is likely due to the fact that the value of α giving the best results for NLRW in G_1 was 0.9995, making NLRW almost identical to NRW.

TABLE III

LEFT: RI FOR ALL 6 DATASETS. HIGHER VALUES OF RI ARE BETTER. EVERY METHOD USED THE PARAMETERS PRODUCING THE BEST RI FOR G_1.
RIGHT: STATISTICAL ANALYSIS FOR RI ON ALL DATASETS. A VALUE OF '1' INDICATES THAT THE ROW IS STATISTICALLY BETTER THAN THE
COLUMN, WHILE A VALUE OF '0' INDICATES THAT THE ROW IS STATISTICALLY WORSE THAN THE COLUMN; A VALUE OF '-' INDICATES
THAT THE ROW AND COLUMN ARE STATISTICALLY INDISTINGUISHABLE. EVERY SUB-ENTRY CORRESPONDS TO THE DIFFERENT
DATASETS: G_1, G_2, W_1, W_2, S, V AND E. (a) RI FOR ALL 6 DATASETS. (b) STATISTICAL ANALYSIS FOR RI

(a)

Dataset	G_1	G_2	W_1	W_2	S	V	E
RW	0.6419	0.8528	0.7938	0.8633	0.8192	0.6518	0.7738
LRW	0.6473	0.7980	0.7954	0.8016	0.7444	0.5285	0.7134
LC	0.6787	0.8260	0.8025	0.8459	0.8111	0.6137	0.7612
RWR	0.6358	0.7809	0.7600	0.7804	0.7169	0.5018	0.6636
NRW	0.6827	0.8698	0.8401	0.8806	0.8502	0.6761	0.8081
NLRW	0.6838	0.8698	0.8385	0.8799	0.8494	0.6752	0.8070

(b)

Method	RW	LRW	LC	RWR	NRW	NLRW
RW	- - - - -	- 1111	- - -1-	-1111	0-000	0 - -00
LRW	-0000	- - - - -	- -000	- - -11	00000	00000
LC	- - -0-	- -111	- - - - -	-1111	- -000	- -000
RWR	-0000	- - -00	-0000	- - - - -	00000	00000
NRW	1-111	11111	- -111	11111	- - - - -	- - - - -
NLRW	1- -11	11111	- -111	11111	- - - - -	- - - - -
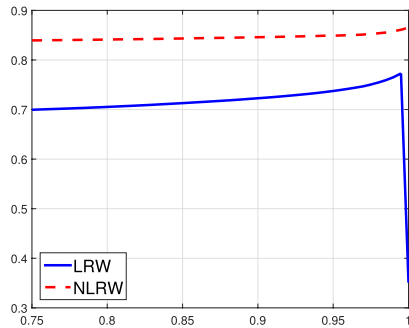
Fig. 8. RI plotted against α. Red: NLRW, Blue: LRW.

Fig. 9. RI plotted against $\frac{1}{\sigma_g}$ for all studied graph-based methods.
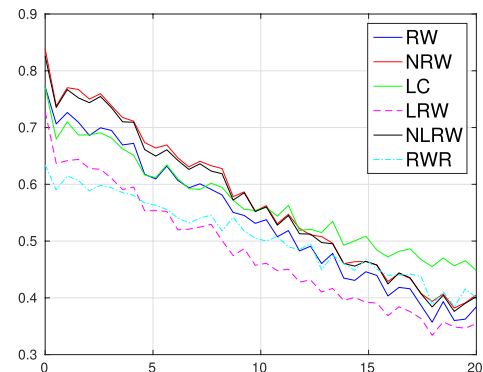

Fig. 10. RI plotted against percentage of misplaced nodes in the first label for all graph-based methods.

Fig. 11. Original images from the Semantic dataset used in the graph-based experiments and randomly generated seeds.

We also carried out a statistical significance analysis of the scores in Table III(a) using an unpaired two sample t-test ($p = 0.05$). We repeated this process for all methods, all datasets and all three evaluation metrics. The results are tabulated in Tables III(b) and IV, which show the statistical superiority of NRW across the different datasets. Regarding the other methods, LC was superior to LRW and RWR but similar to RW except on the VOC12 dataset. NLRW performed similar to NRW given that $\alpha = 0.9995$. We believe that the power of NRW originates from the fact that it directly uses the degree-aware term to control the underlying diffusion schemes. In the arbitrary graph setting, the degree of the nodes captures local similarities and the importance of each node. This is a rich source of information that can greatly improve the graph-segmentation results.

The lazy probability α in NLRW and LRW affects the final clustering result. To study this, we varied α from 0.75 to 1 in steps of 0.05 and measured the RI for both algorithms. As shown in Fig. 8, both methods delivered improved results as α was increased. However, when $\alpha = 1$ the LRW results degraded, hence the LRW algorithm does not degenerate to the RW algorithm when $\alpha = 1$. By contrast, the behavior of NLRW was much smoother, peaking at $\alpha = 1$. In fact, when $\alpha = 1$, the NLRW degenerates to NRW since the node/patient no longer resists changes to its infection status.

The common parameter used by all methods is σ_g used in the construction of the weight matrix. We analyzed the effect of varying this parameter on all methods to better understand how well they can generalize over different parameter choices. We varied $\frac{1}{\sigma_g}$ over the interval $[10, \ldots, 200]$ on the

Semantic100 dataset and report the average RI. As shown in Fig. 9, NRW and NLRW consistently outperformed the other methods for all values of σ_g, showing NRW to be an attractive and reliable method. Fig. 9 shows that for large values of σ_g, the performance of NLRW exceeded that of NRW. It is likely

TABLE IV

STATISTICAL ANALYSIS OF THE GCE AND VOI METRICS ON ALL DATASETS. EVERY METHOD USED THE PARAMETERS PRODUCING THE BEST RI FOR G_1. A VALUE OF '1' INDICATES THAT THE ROW IS STATISTICALLY BETTER THAN THE COLUMN, WHILE A VALUE OF '0' INDICATES THAT THE ROW IS STATISTICALLY WORSE THAN THE COLUMN; A VALUE OF '-' INDICATES THAT THE ROW AND COLUMN ARE STATISTICALLY INDISTINGUISHABLE. EVERY SUB-ENTRY CORRESPONDS TO THE DIFFERENT DATASETS: G_1, G_2, W_1, W_2, S, V AND E

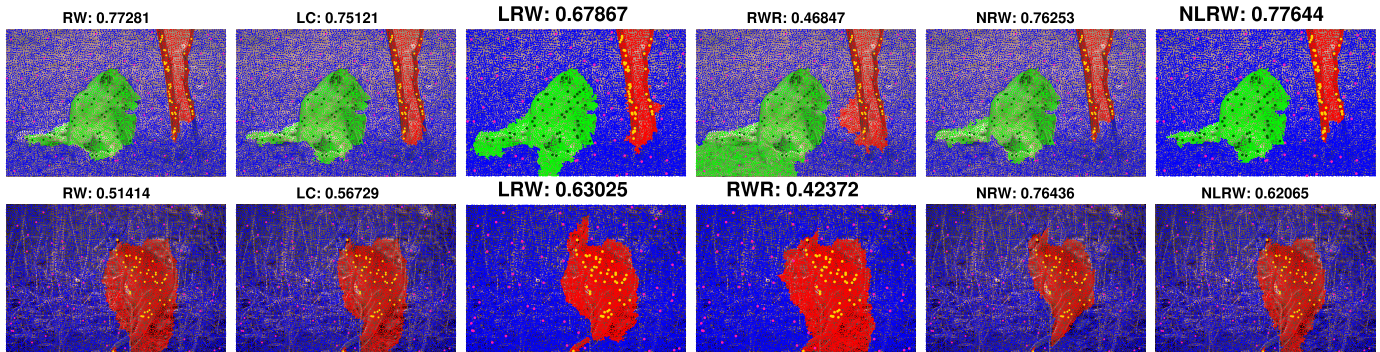| Metric	GCE						VOI					
Method	RW	LRW	LC	RWR	NRW	NLRW	RW	LRW	LC	RWR	NRW	NLRW
RW	- - - - -	- - - 11	- - - 1-	- -111	- - -00	- - -00	- - - - -	-1111	- - -1-	-1111	- - -00	- -00
LRW	- - -00	- - - - -	- - - -0	- - -11	-0000	-0000	-0000	- - - - -	- -000	- - -11	-0000	-0000
LC	- - -0-	- - - -1	- - - - -	- -111	- -000	- -000	- - -0-	- -111	- - - - -	-1111	- -000	- -000
RWR	- -000	- - -00	- -000	- - - - -	00000	00000	- -000	- - -00	-0000	- - - - -	00000	00000
NRW	- - -11	-1111	- -111	11111	- - - - -	- - - - -	- - -11	-1111	- -111	11111	- - - - -	- - - - -
NLRW	- - -11	-1111	- -111	11111	- - - - -	- - - - -	- - -11	-1111	- -111	11111	- - - - -	- - - - -

Fig. 12. Images from the Semantic dataset (seen in Fig. 11) for all methods, $\sigma_g = \frac{1}{90}$, 50 random seeds per class.
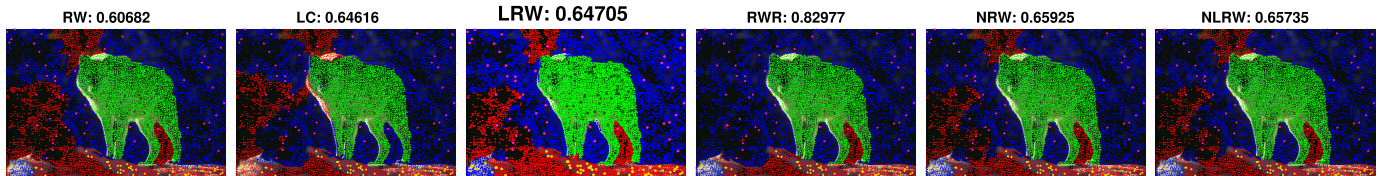

Fig. 13. Effect of mislabeled pixels for 4-infections on the third image from Fig. 11: all methods produce less accurate results. The RWR algorithm delivered more robust results. NRW was the second best method. RI values are above the graph clustering results. The seed colors corresponds to their true label. We do not show mislabeled seeds. Some of the true labels were internally mislabeled as belonging to other regions. The effect of reduced seed quality can be seen in the areas around the wolf. We used the suggested parameters for all methods.

that the presence of the parameter α imparts robustness to NLRW for these values of σ_g. For all methods, the best performing $\frac{1}{\sigma_g}$ value was in the range [40, 50]. Another important factor in seeded image segmentation methods is the seed quality. Manual annotations can be noisy owing to the problem ambiguity or lack of technical expertise or concentration. We studied the effect of varying the seed quality by selecting a fraction of points from the 1st labeled region of each image, which were then randomly applied to another region. These misclassified seeds reflect the low seed quality often present in real world cases. Then, we performed 100 different trials by varying the percentage of misclassified seeds. Fig. 10 shows the mean RI obtained for all methods. The NRW/NLRW algorithm results were better than those of the LC algorithm until about 10% of the seeds were misplaced.

We next examined several exemplar results in detail. Fig. 11 shows the original images used in the experiments with the seeded nodes superimposed as colored dots. Fig. 12 shows the final clustering results across different methods without

any misplaced pixels. Clearly, the performance of NRW was improved while the LRW result suffered from bleeding. NLRW was better than NRW in the first case but worse in the second, and RWR performed the worst in both cases. However, RWR did not always perform the worst compared to other methods, as shown in Fig. 13, where a number of nodes were intentionally mislabeled. Clearly, RWR yielded the best results overall and was able to achieve good results even when some of the seeds were not accurate.

Graph-based approaches can yield significant computational speed-ups since the image dimensionality is reduced to the number of graph nodes. To demonstrate this effect, we selected 30 images from the GrabCut dataset at random and rescaled them to different image sizes (both downscaling and upscaling was performed to cover a wide range of image sizes). When necessary, we cropped the images to ensure that all had the same original size. Then, we ran all the pixel- and graph-based approaches on each scaled image (using random seed points) and plotted the compute times in Fig. 14. All graph-based
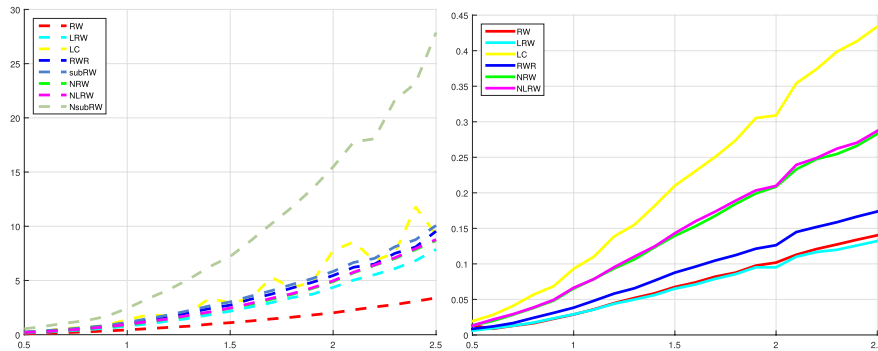
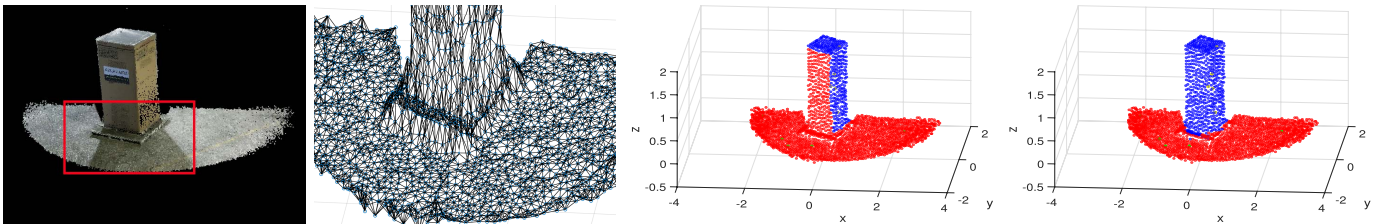Fig. 14. Time in seconds as a function of the number of pixels, left: pixel-based, right: graph-based.

Fig. 15. From left to right: 3D Point cloud captured by Kinect sensor, k-nn graph, k = 8, RW and NRW $\sigma_g = \frac{1}{60}$, $\sigma_h = \frac{1}{10}$, 5 random seeds per class.

approaches consumed considerably less time. For NRW and NLRW, the main overhead originates from computing $\mathbf{D}^{-\frac{1}{2}}$. The large compute time of NsubRW arises from the more complex matrix operations needed. Also, LC was the second slowest method for the pixel case and the slowest in the graph case. LC involves more complex operations on the Laplacian matrix. These results mostly agree with the simulations reported in Table I. We believe that the additional overhead of NRW is less important given its improved performance, especially in the node-based experiments. Clearly, using the RAG allows the modelling of more complex pixel interactions, and also greatly reduces the problem dimension.

X. APPLICATION TO VISUAL POINT CLOUD

To demonstrate the broader promise of this framework, we demonstrate its use on 3D point clouds. Point clouds are now common given the broad use of depth cameras. First, we captured a 3D scene using a Kinect sensor and further processed it using the Point Cloud Library (PCL). We used color and depth information as features to separate objects from ground through a graph-based segmentation task (see Fig. 15). This task can be trivial if one uses only spatial constraints with a prior on where the object should be. We created a k nearest neighbor (k-nn) graph where $k = 8$ and applied both the RW and the NRW algorithms using the extracted features to construct the weight matrix using (6). The seeds were randomly placed around the main areas of the two objects (fridge and ground). The RW method failed to segment the object as it was misled by the shadow. By contrast, NRW was better able to discriminate between the two. We directly applied NRW without modification since it is not limited to regular image grids and can be readily deployed on other graph-related tasks such as video segmentation [45].

XI. CONCLUDING REMARKS

Our proposed models (NRW, NLRW and NsubRW), rely heavily on a node centrality term that drives the underlying graph diffusion processes. This term is an important descriptor of the local interactions that occur between nodes in an image-driven graph. An analogy was made of these interactions as infectious diseases propagating on a graph. NRW expresses the steady state of these diffusion schemes. We analyzed its properties, such as violation of the maximum principle, and provided a theoretical bound on the node degree of pixel-based vs. graph-based methods. We show why NRW is more effective on arbitrary graphs. A lazy random walk variant (NLRW) was also proposed based on the resistance of a node to changes in its infection status. By an extensive experimental evaluation, we showed that NRW is a robust and promising alternative, especially on the graph-based problems, where incorporating the network-aware term is helpful. Future work could potentially use these ideas for other applications where visual information is not represented in a regular way, such as graph spatiotemporal approaches.

APPENDIX

A. Derivation of Iterative NRW Scheme

We now derive (11). Consider the divergence operator $d(.)$ defined in [32]:

$$d(i,t) = -\sum_{i \sim j} \sqrt{w_{ij}} \big(f(i,j,t) - f(j,i,t) \big)$$

where f is a function defined on the edges of the graph (at nodes i and j and time t) and $d(i,t)$ is the net outflow of f at node i and time t. Note that we have defined a dynamic operation $d(i,t)$ rather than $d(i)$. Then, consider the following:

$$I_{i,t+1} = I_{i,t} - d(i,t). \tag{21}$$

When the net outflow is 0, steady state is reached and the infection at node i is no longer changing. In the original RW setting, $f(i, j, t) = \sqrt{w_{ij}} I_{j,t}$ and $f(j, i, t) = \sqrt{w_{ij}} I_{i,t}$ yielding (8). In the case of NRW, normalize f as $f(i, j, t) = \sqrt{w_{ij}} \frac{I_{j,t}}{\sqrt{d_i}\sqrt{d_j}}$ and $f(j, i, t) = \sqrt{w_{ij}} \frac{I_{i,t}}{\sqrt{d_i}\sqrt{d_i}}$. Next, rewrite (21) by using the newly defined $f(i, j, t)$ and $f(j, i, t)$:

$$I_{i,t+1} = I_{i,t} + \sum_{i \sim j} \sqrt{w_{ij}} \big(f(i, j, t) - f(j, i, t) \big)$$

$$= I_{i,t} + \sum_{i \sim j} \frac{w_{ij}}{\sqrt{d_i}} \left(\frac{I_{j,t}}{\sqrt{d_j}} - \frac{I_{i,t}}{\sqrt{d_i}} \right) \qquad (22)$$

which is identical to the NRW iterative scheme in (11).

B. Proofs of the Propositions

1) Proof of Proposition 1:
Proof: Denote by $\mathbf{v}_i$ the ith column of the graph Laplacian $\tilde{\mathbf{L}}$. Using (10) we get:

$$\mathbf{v}_i^\top \mathbf{I}_t = I_{i,t} - \sum_{j \sim i} \frac{w_{ij}}{\sqrt{d_i d_j}} I_{j,t} = \frac{d_i}{d_i} I_{i,t} - \sum_{j \sim i} \frac{w_{ij}}{\sqrt{d_i d_j}} I_{j,t}$$

$$= \sum_{j \sim i} \frac{w_{ij}}{\sqrt{d_i}} \left(\frac{I_{i,t}}{\sqrt{d_i}} - \frac{I_{j,t}}{\sqrt{d_j}} \right) = -\Delta I_{i,t}.$$

Stacking the $\Delta I_{i,t}$ yields:

$$\mathbf{\Delta I}_t = \begin{pmatrix} \Delta I_{1,t} \\ \vdots \\ \Delta I_{N,t} \end{pmatrix} = - \begin{pmatrix} \mathbf{v}_1^\top \mathbf{I}_t \\ \vdots \\ \mathbf{v}_N^\top \mathbf{I}_t \end{pmatrix} = - \begin{pmatrix} \mathbf{v}_1^\top \\ \vdots \\ \mathbf{v}_N^\top \end{pmatrix} \mathbf{I}_t$$

which gives (12). Further, note that $J_n(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \tilde{\mathbf{L}} \mathbf{x}$. Replace $\mathbf{x}$ by the infection probabilities $\mathbf{I}_t$. Reorder and partition $\tilde{\mathbf{L}}$, $\mathbf{\Delta I}_t$ and $\mathbf{I}_t$ to obtain rows corresponding to un-labeled (u) and labeled (l) nodes, i.e

$$\tilde{\mathbf{L}} = \begin{pmatrix} \tilde{\mathbf{L}}_u \\ \tilde{\mathbf{L}}_l \end{pmatrix}, \quad \mathbf{\Delta I}_t = \begin{pmatrix} \mathbf{\Delta I}_{u,t} \\ \mathbf{\Delta I}_{l,t} \end{pmatrix}, \quad \mathbf{I}_t = \begin{pmatrix} \mathbf{I}_{u,t} \\ \mathbf{I}_{l,t} \end{pmatrix} \qquad (23)$$

Then re-order the columns of $\tilde{\mathbf{L}}$ so that:

$$\tilde{\mathbf{L}} = \begin{pmatrix} \tilde{\mathbf{L}}_{uu} & \tilde{\mathbf{L}}_{ul} \\ \tilde{\mathbf{L}}_{lu} & \tilde{\mathbf{L}}_{ll} \end{pmatrix}. \qquad (24)$$

Since $\tilde{\mathbf{L}}$ is positive semi-definite, apply a process similar to [7] to minimize $J_n(\mathbf{I_t})$:

$$\tilde{\mathbf{L}}_{uu} \mathbf{I}_{u,t} = -\tilde{\mathbf{L}}_{ul} \mathbf{I}_{l,0}. \qquad (25)$$

Clearly $\mathbf{I}_{l,t} = \mathbf{I}_{l,0} \; \forall t$. Meanwhile, using (12) yields:

$$\mathbf{\Delta I}_t = -\tilde{\mathbf{L}} \mathbf{I}_t \Rightarrow \mathbf{\Delta I}_{u,t} = -\tilde{\mathbf{L}}_u \mathbf{I}_t \qquad (26)$$

$$= -\tilde{\mathbf{L}}_{uu} \mathbf{I}_{u,t} - \tilde{\mathbf{L}}_{ul} \mathbf{I}_{l,t} \qquad (27)$$

For $t \to \infty \Rightarrow \mathbf{\Delta I}_{u,t} \to 0$ hence using (27) yields (25). $\qquad \square$

2) Proof of Proposition 2:
Proof: When $t \to t_\infty$ then:

$$\Delta I_{i,t_\infty} = 0 \overset{(11)}{\Longrightarrow} \sum_{j \sim i} \frac{w_{ij}}{\sqrt{d_i}} \left(\frac{I_{j,t_\infty}}{\sqrt{d_j}} - \frac{I_{i,t_\infty}}{\sqrt{d_i}} \right) = 0$$

$$\Rightarrow \sum_{j \sim i} \frac{w_{ij}}{\sqrt{d_i}} \frac{I_{j,t_\infty}}{\sqrt{d_j}} = \sum_{j \sim i} \frac{w_{ij}}{\sqrt{d_i}} \frac{I_{i,t_\infty}}{\sqrt{d_i}} = \sum_{j \sim i} \frac{w_{ij}}{d_i} I_{i,t_\infty}$$

$$= \frac{I_{i,t_\infty}}{d_i} \sum_{j \sim i} w_{ij} = I_{i,t_\infty}. \qquad \square$$

3) Proof of Proposition 3:
Proof: Suppose node i has at least one neighbor j_0. Then:

$$I_{i,t_\infty} = \frac{1}{\sqrt{d_i}} \sum_{j \sim i} \frac{w_{ij}}{\sqrt{d_j}} I_{j,t_\infty} = \frac{1}{\sqrt{d_i}} \frac{w_{ij_0}}{\sqrt{d_{j_0}}} I_{j_0,t_\infty} + \epsilon$$

for $d_{j_0} > 0$, $d_i > 0$ and $\epsilon \geq 0$. Then, given some λ : $|\lambda| < 1$, $\lambda \neq 0$, pick $d_{j_0} = \frac{1}{d_i} w_{ij_0}^2 I_{j_0,t_\infty}^2 \lambda^2$ which implies that $I_{i,t_\infty} = \frac{1}{|\lambda|} + \epsilon \geq 1$. Since the maximum principle is violated, I_{i,t_∞} loosely represents some infection's probability hence the sum of I_{i,t_∞} is $\neq 1$. Therefore, unlike [7], all M systems of linear equations need now be solved. $\qquad \square$

Proposition 4: It is possible to quantify the difference between using a pixel-based scheme (4-neighbors) vs. using a node-based approach where an arbitrary number of neighbors is present. Consider the quantity $E\big[|d_i^g - d_i^p|\big]$ where the random variables d_i^g and d_i^p denote the degree of a node i using the graph or pixel scheme respectively. Let image I have pixel values $\in [0, \; m-1]$. Then, $E\big[|d_i^g - d_i^p|\big]$ is lower bounded by $f(m, \sigma_g) E\big[|C-4|\big]$, where C is a random variable representing the number of neighbors for every node in the graph and $f(m, \sigma_g)$ is a function of m and σ_g.

Proof: Suppose the corresponding weights w_{ij}^g and w_{ij}^p are equal and given by (6) with $\sigma_h = 0$. Typically, the number C of neighbors is independent of the similarities between nodes based on color since node connections are based either on distance e.g. by using a k-nearest neighbor graph (k-nn) or a watershed topology (RAG). Hence

$$E\big[|d_i^g - d_i^p|\big] = E\Big[|\sum_{j \sim i} w_{ij}^g - \sum_{j=1}^4 w_{ij}^p|\Big]$$

$$= E\Big[\sum_{j=1}^{|C-4|} w_{ij}^g\Big] = E\big[|C - 4|\big] E[w_{ij}]$$

where $E[w_{ij}]$ is the expected value of the similarity between any two connected nodes i and j. Images are spatially correlated, hence adding descriptors of the image structure should increase the similarity between any two nodes. Consider a completely random image where $g_i \sim \mathcal{U}[0, \; m-1]$, where $m-1$ is the maximum luminance. The absolute difference of two iid uniform random variables g_i, g_j has the following pdf:

$$f_X(x) = \begin{cases} -\dfrac{2}{(m-1)^2} x + \dfrac{2}{m-1}, & \text{if } x \in [0, \; m-1] \\ 0, & \text{else.} \end{cases} \qquad (28)$$

Then, we obtain:

$$E[w_{ij}] = \int_0^{m-1} \left(\frac{2}{m-1} - \frac{2x}{(m-1)^2} \right) e^{-\left(\frac{x}{\sigma_g}\right)^2} dx$$

$$= \frac{e^{-\left(\frac{m-1}{\sigma_g}\right)^2} + \sqrt{\pi}\frac{m-1}{\sigma_g}\operatorname{erf}\left(\frac{m-1}{\sigma_g}\right) - 1}{\left(\frac{m-1}{\sigma_g}\right)^2} = f(m, \sigma_g)$$

where $\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}}\int_0^z e^{-t^2} dt$. For $m = 256$ and $\sigma_g = \frac{1}{90}$ we obtain: $E[w_{ij}] \approx 7.72 * 10^{-5}$. Real world images have structural correlations and as a result $E[w_{ij}]$ is lower bounded by $f(m, \sigma_g)$. The expected value $E\big[|C - 4|\big]$ is not trivial to compute, since for different graph structures there can be different node degree distributions. The simplest case occurs using a k-nn graph which degenerates the random variable C to a constant k, whence $E\big[|C - 4|\big] = |k - 4|$. The resulting bound is not very tight, but the experimental section demonstrates the merits of the node-based NRW against the pixel version.

C. Extending the SubRW Framework to Include the Smoothness Term of NRW

This section is a high level analysis of subRW [14] necessary to demonstrate the use of the NRW smoothness term in more complex RW variants. We refer the reader to [14] for details. Consider the following optimization scheme (NsubRW):

$$\bar{O}^{l_k}$$

$$= \frac{1}{2}\sum_{i \sim j} w_{ij}\left(\frac{\bar{r}_{im}^{l_k}}{\sqrt{d_i}} - \frac{\bar{r}_{jm}^{l_k}}{\sqrt{d_j}}\right)^2 + \frac{1}{2}\sum_{i=1}^{N}\frac{(d_i + \lambda g_i)c_i}{d_i(1 - c_i)}\left(\bar{r}_{im}^{l_k} - b_{im}^{l_k}\right)^2$$

$$+ \frac{1}{2}\sum_{i=1}^{N}\frac{\lambda}{d_i}u_i^k\left(\bar{r}_{im}^{l_k} - 1\right)^2 + \frac{1}{2}\sum_{t=1, t\neq k}^{K}\sum_{i=1}^{N}\frac{\lambda}{d_i}u_i^t\bar{r}_{im}^{l_k^2} \qquad (29)$$

which re-expresses the subRW scheme by taking into account d_i and d_j. If $\lambda = 0$, $c_i = 0\ \forall i$, yielding $J_n(\mathbf{x})$ which is minimized by NRW. This shows an interesting dichotomy: while subRW reduces to RW, NsubRW reduces to NRW for $\lambda = 0$ and $c_i = 0\ \forall i$. Following the same process as in [14], vectorize (29) and zero the partial derivative with respect to $\mathbf{r}_m^{-l_k} = [\bar{r}_{im}^{l_k}]\ \forall i$. It then follows that the solution of NsubRW is given by: $\bar{\mathbf{r}}_m^{l_k} = \bar{\mathbf{E}}^{-1}((\mathbf{I} - \mathbf{D_c})\bar{\mathbf{u}}^k + \mathbf{D_c}\mathbf{b}_m^{l_k})$, where $\bar{\mathbf{E}} = \mathbf{I} - \mathbf{D_a}^{-1}\mathbf{D}_\eta\mathbf{D}\mathbf{S}$, $\mathbf{S} = \mathbf{D}^{-\frac{1}{2}}\mathbf{W}\mathbf{D}^{-\frac{1}{2}}$, $\mathbf{D}$ is the $N \times N$ degree matrix, $\mathbf{W}$ is the $N \times N$ weight matrix, N is the number of pixels in the image, $\mathbf{I}$ is a $N \times N$ diagonal matrix of ones, $\mathbf{D_c}$ is a $N \times N$ diagonal matrix containing the restarting probabilities c_i, $\bar{\mathbf{u}}^k$ is a $N \times 1$ vector related to the GMM prior, $\mathbf{b}_m^{l_k}$ is a $N \times 1$ vector capturing the seed information, $\mathbf{D_g}$ is a $N \times N$ diagonal matrix also related to the label prior and $\mathbf{D}_a = \mathbf{D} + \lambda\mathbf{D_g}$. We can also consider schemes related to RWR:

$$\bar{O}^{l_k} = \frac{1}{2}\sum_{i \sim j} w_{ij}\left(\frac{\bar{r}_{im}^{l_k}}{\sqrt{d_i}} - \frac{\bar{r}_{jm}^{l_k}}{\sqrt{d_j}}\right)^2 + \frac{1}{2}\sum_{i=1}^{N}\frac{c}{1-c}\left(\bar{r}_{im}^{l_k} - b_{im}^{l_k}\right)^2$$

where $c_i = c$, $\lambda = 0$ in (29). The prior process then yields: $\bar{\mathbf{r}}_m^{l_k} = \bar{\mathbf{E}}^{-1}\mathbf{D_c}\mathbf{b}_m^{l_k}$, where $\bar{\mathbf{E}} = \mathbf{I} - (1 - c)\mathbf{S}$. We refer to this scheme as NRWR. This solution is equivalent to LRW where the lazy probability α is set to $1 - c$. This demonstrates how RWR, NRWR and LRW are related to each other. In our graph-based experiments, the NRWR performed much better than RWR but its performance was upper bounded by LRW, reaching a peak value when $\alpha = 1 - c$.

REFERENCES

[1] P. Arbeláez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 5, pp. 898–916, May 2011.

[2] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vis.*, vol. 1, no. 4, pp. 321–331, 1988.

[3] M. Clark and A. C. Bovik, "Experiments in segmenting texton patterns using localized spatial filters," *Pattern Recognit.*, vol. 22, no. 6, pp. 707–717, 1989.

[4] J. Fan, D. K. Y. Yau, A. K. Elmagarmid, and W. G. Aref, "Automatic image segmentation by integrating color-edge extraction and seeded region growing," *IEEE Trans. Image Process.*, vol. 10, no. 10, pp. 1454–1466, Oct. 2001.

[5] P. Kohli, A. Osokin, and S. Jegelka, "A principled deep random field model for image segmentation," in *Proc. IEEE CVPR*, Jun. 2013, pp. 1971–1978.

[6] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 888–905, Aug. 2000.

[7] L. Grady, "Random walks for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 11, pp. 1768–1783, Nov. 2006.

[8] J. Zhang, J. Zheng, and J. Cai, "A diffusion approach to seeded image segmentation," in *Proc. IEEE Conf. Comp. Vis. Pat. Recogn. (CVPR)*, Jun. 2010, pp. 2125–2132.

[9] B. Ham, D. Min, and K. Sohn, "A generalized random walk with restart and its application in depth up-sampling and interactive segmentation," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2574–2588, Jul. 2013.

[10] T. H. Kim, K. M. Lee, and S. U. Lee, "Generative image segmentation using random walks with restart," in *Proc. ECCV*, Oct. . 2008, pp. 264–275.

[11] J. Shen, Y. Du, W. Wang, and X. Li, "Lazy random walks for superpixel segmentation," *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1451–1462, Apr. 2014.

[12] Y. Liang, J. Shen, X. Dong, H. Sun, and X. Li, "Video supervoxels using partially absorbing random walks," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 5, pp. 928–938, May 2016.

[13] L. Grady, "Multilabel random walker image segmentation using prior models," in *Proc. IEEE Conf. Comput. Vis. Pat. Recogn. (CVPR)*, vol. 1. Jun. 2005, pp. 763–770.

[14] X. Dong, J. Shen, L. Shao, and L. Van Gool, "Sub-Markov random walk for image segmentation," *IEEE Trans. Image Process.*, vol. 25, no. 2, pp. 516–527, Feb. 2016.

[15] X. Dong, J. Shen, and L. Van Gool, "Segmentation using subMarkov random walk," in *Energy Minimization Methods in Computer Vision and Pattern Recognition*. Springer, 2015, pp. 237–248.

[16] W. Casaca, L. G. Nonato, and G. Taubin, "Laplacian coordinates for seeded image segmentation," in *Proc. IEEE Conf. Comp. Vis. Pat. Recogn. (CVPR)*, Jun. 2014, pp. 384–391.

[17] W. Casaca, A. Paiva, and L. G. Nonato, "Spectral segmentation using cartoon-texture decomposition and inner product-based metric," in *Proc. 24th Conf. Grap., Patt. Imag.*, Aug. 2011, pp. 266–273.

[18] W. Casaca, "Graph laplacian for spectral clustering and seeded image segmentation," Ph.D. dissertation, Dept. ICMC-USP, Univ. São Paulo, São Paulo, Brazil, 2014.

[19] J. Shen, Y. Du, and X. Li, "Interactive segmentation using constrained Laplacian optimization," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 7, pp. 1088–1100, Jul. 2014.

[20] W. Yang, J. Cai, J. Zheng, and J. Luo, "User-friendly interactive image segmentation through unified combinatorial user inputs," *IEEE Trans. Image Process.*, vol. 19, no. 9, pp. 2470–2479, Sep. 2010.

[21] J. Bai and X. Wu, "Error-tolerant scribbles based interactive image segmentation," in *Proc. IEEE Conf. Comp. Vis. Pat. Recogn.*, Jun. 2014, pp. 392–399.

[22] C. Couprie, L. Grady, L. Najman, and H. Talbot, "Power watershed: A unifying graph-based optimization framework," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 7, pp. 1384–1399, Jul. 2011.

[23] B. Wang and Z. Tu, "Affinity learning via self-diffusion for image segmentation and clustering," in *Proc. CVPR*, Jun. 2012, pp. 2312–2319.

[24] L. C. Freeman, "Centrality in social networks conceptual clarification," *Social Netw.*, vol. 1, no. 3, pp. 215–239, 1978.

[25] C. G. Bampis and P. Maragos, "Unifying the random walker algorithm and the SIR model for graph clustering and image segmentation," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2015, pp. 2265–2269.

[26] H. W. Hethcote, "The mathematics of infectious diseases," *SIAM Rev.*, vol. 42, no. 4, pp. 599–653, 2000.

[27] E. B. Postnikov and I. M. Sokolov, "Continuum description of a contact infection spread in a SIR model," *Math. Biosci.*, vol. 208, no. 1, pp. 205–215, 2007.

[28] P. G. Doyle and J. L. Snell, *Random Walks and Electric Networks* (Carus Mathematical Monographs). Washington, DC, USA: Mathematical Association of America, 1984.

[29] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, vol. 2. Hoboken, NJ, USA: Wiley, 1989.

[30] X. Zhu, Z. Ghahramani, and J. Lafferty, "Semi-supervised learning using Gaussian fields and harmonic functions," in *Proc. 12th ICML*, vol. 3. 2003, pp. 912–919.

[31] F. Wang and C. Zhang, "Label propagation through linear neighborhoods," *IEEE Trans. Knowl. Data Eng.*, vol. 20, no. 1, pp. 55–67, Jan. 2008.

[32] A. Elmoataz, O. Lézoray, and S. Bougleux, "Nonlocal discrete regularization on weighted graphs: A framework for image and manifold processing," *IEEE Trans. Image Process.*, vol. 17, no. 7, pp. 1047–1060, Jul. 2008.

[33] Y. Artan and I. S. Yetik, "Improved random walker algorithm for image segmentation," in *Proc. IEEE Southwest Symp. Image Anal. Interpretation*, May 2010, pp. 89–92.

[34] D. Zhou, O. Bousquet, T. N. Lal, J. Weston, and B. Schölkopf, "Learning with local and global consistency," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 16. 2004, pp. 321–328.

[35] A. Trémeau and P. Colantoni, "Regions adjacency graph applied to color image segmentation," *IEEE Trans. Image Process.*, vol. 9, no. 4, pp. 735–744, Apr. 2000.

[36] F. Meyer, "Topographic distance and watershed lines," *Signal Process.*, vol. 38, no. 1, pp. 113–125, 1994.

[37] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012.

[38] F. R. Chung, *Spectral Graph Theory* (Regional Conference Series in Mathematics). Providence, RI, USA: American Mathematical Society, 1997.

[39] C. Rother, V. Kolmogorov, and A. Blake, "'GrabCut': Interactive foreground extraction using iterated graph cuts," *ACM Trans. Graph.*, vol. 23, no. 3, pp. 309–314, Aug. 2004.

[40] F. Andrade and E. V. Carrera, "Supervised evaluation of seed-based interactive image segmentation algorithms," in *Proc. 20th Symp. Signal Process., Images Comput. Vis.*, Sep. 2015, pp. 1–7.

[41] S. Alpert, M. Galun, R. Basri, and A. Brandt, "Image segmentation by probabilistic bottom-up aggregation and cue integration," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2007, pp. 1–8.

[42] H. Li, J. Cai, T. N. A. Nguyen, and J. Zheng, "A benchmark for semantic image segmentation," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2013, pp. 1–6.

[43] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The Pascal visual object classes (VOC) challenge," *Int. J. Comput. Vis.*, vol. 88, no. 2, pp. 303–338, Sep. 2009.

[44] J. Shi, Q. Yan, L. Xu, and J. Jia, "Hierarchical image saliency detection on extended CSSD," *IEEE Trans. Pat. Anal. Mach. Intell.*, vol. 38, no. 4, pp. 717–729, Apr. 2016.

[45] H. Jiang, G. Zhang, H. Wang, and H. Bao, "Spatio-temporal video segmentation of static scenes and its applications," *IEEE Trans. Multimedia*, vol. 17, no. 1, pp. 3–15, Jan. 2015.

Christos G. Bampis (SM'15) received the M.Eng. Diploma degree in electrical engineering from the National Technical University of Athens in 2014. He is currently pursuing the Ph.D. degree with the Laboratory for Image and Video Engineering, The University of Texas at Austin.

His research interests include image and video quality assessment, and computer vision with an emphasis on graph-theoretic approaches for image segmentation.

Petros Maragos (F'95) received the M.Eng. Diploma degree in electrical engineering from the National Technical University of Athens (NTUA) in 1980 and the M.Sc. and Ph.D. degrees from Georgia Tech, Atlanta, in 1982 and 1985, respectively. In 1985, he joined the Faculty of the Division of Applied Sciences, Harvard University, Boston, MA, USA, where he was a Professor of Electrical Engineering, affiliated with the Harvard Robotics Laboratory, for eight years. In 1993, he joined the Faculty of the School of ECE, Georgia Tech, affiliated with its Center for Signal and Image Processing. From 1996 to 1998 he had a joint appointment as the Director of research with the Institute of Language and Speech Processing in Athens. He has held Visiting Scientist positions with MIT in fall 2012 and the University of Pennsylvania in fall 2016. Since 1999, he has been a Professor with the NTUA School of ECE, where he is currently the Director of the Intelligent Robotics and Automation Laboratory. He has authored numerous papers, book chapters, and has also co-edited three Springer research books, one on multimodal processing and two on shape analysis. His research and teaching interests include signal processing, systems theory, machine learning, image processing and computer vision, audio and speech/language processing, cognitive systems, and robotics. He has served as a member of the IEEE SPS committees on DSP, IMDSP, and MMSP. He has also served as a member of the Greek National Council for Research and Technology. He has served as an Associate Editor of the IEEE TRANSACTIONS ON ASSP and the IEEE TRANSACTIONS ON PAMI. He was an Editorial Board Member and a Guest Editor of several journals on signal processing, image analysis, and vision. He was a Co-Organizer of several conferences and workshops, including VCIP'92 (GC), ISMM'96 (GC), MMSP'07 (GC), ECCV'10 (PC), ECCV'10 Workshop on Sign, Gesture and Activity, 2011 & 2014 Dagstuhl Symposia on Shape, IROS'15 Workshop on Cognitive Mobility Assistance Robots, and EUSIPCO 2017 (GC).

He is a recipient or co-recipient of several awards for his academic work, including a 1987-1992 US NSF Presidential Young Investigator Award, the 1988 IEEE ASSP Young Author Best Paper Award, the 1994 IEEE SPS Senior Best Paper Award, the 1995 IEEE W.R.G. Baker Prize for the most outstanding original paper, the 1996 Pattern Recognition Society's Honorable Mention best paper award, the best paper award from the CVPR-2011 Workshop on Gesture Recognition. He received the 2007 EURASIP Technical Achievements Award for contributions to nonlinear signal processing, systems theory, image and speech processing. In 2010 he was elected Fellow of EURASIP for his research contributions. He has been elected as the IEEE SPS Distinguished Lecturer from 2017 to 2018.

Alan Conrad Bovik (F'96) holds the Cockrell Family Endowed Regents Chair in engineering with The University of Texas at Austin, where he is currently the Director of the Laboratory for Image and Video Engineering, the Department of Electrical and Computer Engineering, the Institute for Neuroscience. He has authored over 800 technical articles in these areas and holds several U.S. patents. His publications have been cited over 50,000 times in the literature, his current h-index is about 85, and he is listed as a Highly-Cited Researcher by Thompson Reuters. His Erdös number is four. His several books include the companion volumes The Essential Guides to Image and Video Processing (Academic Press, 2009). His research interests include image and video processing, digital television and digital cinema, computational vision, and visual perception.

Dr. Bovik received the Primetime Emmy Award for Outstanding Achievement in Engineering Development from the Television Academy in 2015, for his work on the development of video quality prediction models which have become standard tools in broadcast and post-production houses throughout the television industry. He has also received a number of major awards from the IEEE Signal Processing Society, including the Society Award, the Technical Achievement Award in 2005, the Best Paper Award in 2009, the Signal Processing Magazine Best Paper Award in 2013, the Education Award in 2007, the Meritorious Service Award in 1998, and (coauthor) the Young Author Best Paper Award in 2013. He also received the IEEE Circuits and Systems for Video Technology Best Paper Award in 2016. He also was named recipient of the Honorary Member Award of the Society for Imaging Science and Technology for 2013. He received the SPIE Technology Achievement Award for 2012. He was the IS&T/SPIE Imaging Scientist of the Year for 2011. He is also a recipient of the Joe J. King Professional Engineering Achievement Award in 2015 and the Hocott Award for Distinguished Engineering Research in 2008, from the Cockrell School of Engineering, The University of Texas at Austin. He received the Distinguished Alumni Award from the University of Illinois at Champaign–Urbana in 2008. He is a fellow of the Optical Society of America and the Society of Photo-Optical and Instrumentation Engineers. He is a member of the Television Academy, the National Academy of Television Arts and Sciences and the Royal Society of Photography.

Professor Bovik also co-founded and was the longest-serving Editor-in-Chief of the IEEE TRANSACTIONS ON IMAGE PROCESSING from 1996 to 2002, He created and served as the First General Chair of the IEEE International Conference on Image Processing, held in Austin, Texas, in 1994. His many other professional society activities include the Board of Governors, the IEEE Signal Processing Society from 1996 to 1998; the Editorial Board, the Proceedings of the IEEE from 1998 to 2004; and a Series Editor for Image, Video, and Multimedia Processing, Morgan and Claypool Publishing Company since 2003. He was also the General Chair of the 2014 Texas Wireless Symposium, held in Austin in 2014.

Dr. Bovik is a registered Professional Engineer with the State of Texas and is a Frequent Consultant to legal, industrial and academic institutions.