

## GRID MACROSCOPIC GROWTH LAW AND ITS APPLICATION TO IMAGE INFERENCE

BY

NATALIYA PORTMAN (*Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada*),

ULF GREANDER (*Division of Applied Mathematics, Brown University, Providence, Rhode Island, 02912*),

AND

EDWARD R. VRSCAY (*Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada*)

**Abstract.** In computational anatomy large deformations of anatomical structures observed in images are represented by diffeomorphic flows on the background space of coordinates. They are usually *maximum a posteriori* (MAP) estimates obtained by minimization of the cost function (posterior energy) consistent with the material properties of an organism. This paper constructs the underlying transformations induced by biological growth according to the Growth as Random Iterated Diffeomorphisms (GRID) model proposed by U. Grenander. They are diffeomorphic flows generated by the GRID macroscopic growth integro-differential equation that emphasizes dependency of the flow on such GRID variables as the Poisson intensity of cell decisions and relative rate of expansion/contraction.

We explore some cost function models that yield biologically meaningful estimates of these growth parameters. Namely, we seek a prior energy that measures cell activities represented by the Poisson intensity function. Using the macroscopic growth law we formulate an optimal control problem where the GRID variables are optimal controls that force an image of the initial organism to be continuously transformed into an image of the grown organism.

---

Received May 7, 2009.

2000 *Mathematics Subject Classification.* Primary 62M40, 62F10, 62F15; Secondary 49N45, 65K05, 92C15.

N. Portman was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) in the form of a Postgraduate Scholarship and by the Ontario Ministry of Training, Colleges and Universities in the form of an Ontario Graduate Scholarship.

E. R. Vrsnay was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) in the form of a Discovery Grant.

*E-mail address:* nataliya.portman@mcgill.ca

*E-mail address:* ulf.grenander@gmail.com

*E-mail address:* ervrsnay@math.uwaterloo.ca

We apply the Polak-Ribière conjugate gradient algorithm for direct estimation of the growth parameters from given images. Then the biological mapping is automatically obtained from estimated growth parameters. The accuracy of GRID variable and image estimates obtained by the inference algorithm depends on the value of the weighting coefficient of the prior energy. We propose an experimental evaluation of this coefficient and reveal growth patterns expressed in GRID variables hidden in confocal micrographs of Wingless gene expression patterns in the larval *Drosophila* wing disc.

## I. Introduction

This paper is concerned with the application of the Growth as Random Iterated Diffeomorphisms (GRID) model introduced by U. Grenander [10, 11] to direct image inference of the growth properties of an organism. We have been particularly inspired by the book “Endless Forms Most Beautiful” by Sean B. Carroll, which provides a new look at biological development via visualizing body building gene expression patterns. It may be argued that the GRID approach is particularly well-suited for the modeling of biological growth in terms of gene expression patterns. We shall support this claim by means of a GRID-based analysis of micrographs of Wingless expression patterns in the *Drosophila* wing disc.

Specifically, the isotropic version of the two-dimensional GRID model is considered that allows for approximation of growth-related deformation on a large time scale using a “thermodynamic limit” equation referred to as a macroscopic growth law. According to this deterministic equation the diffeomorphic transformation is generated and controlled by the Poisson intensity of elementary cell events (cell divisions, for instance).

The cost function in the form of the posterior probability is implemented as the optimality criterion for the unobserved Poisson intensity, one of the GRID growth parameters, which is subject to the macroscopic growth equation at any instant of time. As a biological adjustment of the cost function two prior probability models of the Poisson intensity field are discussed. These are a Gaussian random field and a sample-function density for a doubly stochastic Poisson point process. It is shown that the latter model is more general, yielding more realistic growth patterns.

*Diffeomorphic transformations for shape analysis and their estimation methods.* One of the major tasks of computational anatomy [1] is the inference of diffeomorphic transformations on the compact background space of coordinates underlying shape changes of human anatomy as seen in images. Diffeomorphic mappings provide the precise description of morphological differences between anatomical structures necessary for analysis and detection of abnormal development. Estimation techniques for the mappings that have been proposed in the literature are based on various cost function models.

A smooth mapping of anatomical and geometric landmarks onto their homologues from specimen to specimen was pioneered by F. Bookstein [2, 3, 4]. Such a transformation is modeled as a thin plate spline that interpolates the distances between the corresponding landmarks and minimizes the bending energy of any such interpolating function. The cost function is the bending energy defined as the sum of squares of all the second derivatives of the mapping.

In the elastic matching algorithm [5, 6] the cost function is the squared  $L^2$  distance between major anatomic features (sulci and gyri of the cortical surface) in one image and their homologues in another image. The elastic transformation for image registration is obtained by solving the equations governing the deformation of an inhomogeneous elastic object where the external force attempts to minimize the cost function.

Such transformations allow only for small deformations, and for large deformations diffeomorphic flows have been used. They are generated by continuum mechanics equations of motion and are usually constrained to the set of transformations consistent with the material properties of brain anatomy under study. Algorithms for the estimation of large deformations are based on a Navier-Stokes fluid model [7, 8, 9]. A viscous fluid flow law allows for large, nonlinear deformations while maintaining a smooth transformation of a template image. In the Bayesian view a viscous transformation that registers the template image with an image under study corresponds to a maximization of the posterior distribution. Then the cost function is the sum of the likelihood and prior probability density functions. The likelihood density function models the image under study as a sample from a Gaussian process with the mean given by the template image under a viscous transformation. The prior probability density gives large probability to viscous fluid transformations. It is given in Gibbs form with the Gibbs potential defined as the Lagrangian energy density for a viscous fluid flow.

*Previous work on development and applications of the GRID model to image inference.* The Growth as Random Iterated Diffeomorphisms (GRID) model [10, 11] introduced by U. Grenander constructs complete shape transformations according to fundamental biological principles of growth. The GRID model offers a characterization of growth seen in images in terms of biologically meaningful GRID variables such as locations of elementary cell decisions (e.g., cell division, cell enlargement, cell death, etc.) called *seeds*, local deformation patterns and Poisson intensities underlying seed placements.

Growth deformation on a small time scale is a composition of several, elementary, local deformations each resulting from a primitive cell decision, a highly localized event modeled as a point of the inhomogeneous time-space Poisson process (*seed*). Depending on the nature of a seed, local deformation patterns can represent unipolar (growth or decay in one preferential direction) or bipolar (bi-directional growth or decay) development. It can be shown that a continuous space-time approximation of the GRID model is a stochastic differential equation representing the growth pattern as a Poisson-driven Markov process.

Growth deformation on a large time scale underlying visible shape changes is due to a large number of primitive cell decisions. The macroscopic growth pattern can be decomposed into elementary deformations representing locally isotropic (angle-independent) growth or decay. It is approximated by the solution to a macroscopic growth integro-differential equation known as a ‘thermodynamic limit’ equation [11]. This is a deterministic differential equation that generates a biological deformation as a result of an infinite number of elementary biological events at each instant of time. It represents typical growth averaged over all seed contributions.

The goal of this model is to reveal a hidden picture of biological growth via image inference of GRID parameters. A method for their estimation has been proposed in

[12]. First, a transformation minimizing the cost function that measures the mismatch between the deformed and target images and their boundaries is estimated. Second, an observation model is imposed on the estimated transformation in the form of the growth-induced deformation prescribed by the GRID model with additive Gaussian noise. Then the GRID variables are inferred from the observed deformation using a maximum-likelihood approach.

Since the GRID parameters define the growth-induced transformation, it is preferable to estimate them directly from image data. As such, we have developed an algorithm for direct image inference of the GRID variables [13]. It is based on a simpler, isotropic version of the GRID model that allows us to reduce the complexity of mathematical modeling and computational effort involved in the estimation of the growth parameters presented in [12]. In [13] we employ the macroscopic growth law to construct a biological transformation that depends on the Poisson intensity of cell decisions and the relative rate of expansion/contraction. Considering biological growth with the uniform intensity of cell divisions, or, in other words, assuming that seeds are distributed according to a homogeneous Poisson point process in the organism's domain, we specify the cost function in the form of the Gaussian log-likelihood of the relative rate of growth/decay. Then the resulting unconstrained optimal control problem is solved on the evolving Darcyan<sup>1</sup> coordinate system of a growing organism's domain by implementing the Polak-Ribière version of the conjugate gradient method.

Using image data from the larval development of the *Drosophila* wing disc we have obtained biologically meaningful estimates of the growth-induced deformation. However, the proposed likelihood-based estimation method has some shortcomings. It cannot generally be run to full convergence since the diffeomorphic property of the transformation is not preserved. Also, the estimated transformation registers only image subregions with the highest mismatch of grey levels.

*Connection of the research work presented in this paper with image analysis of gene expression patterns.* In this paper we build upon and extend our previous work on direct image inference of GRID growth variables described in [12]. We are motivated by the recent advances in developmental biology described by S. B. Carroll in [15] that shed light on the role that gene expression patterns play in the positioning and shaping of future adult structures of a developing organism. As such, our efforts are focused on a biological adjustment of the original cost function model by including a prior energy function that measures cell activities that are naturally represented by the Poisson parameter of intensity of elementary cell decisions (cell divisions, cell enlargements, cell movements, cell death, etc.). Such a cost function will find new potential applications in the analysis of micrographs of gene expression patterns in embryos. If there is experimental evidence that the observed dynamics of levels of a particular gene expression (or concentrations of a gene product) depends on an underlying biological process of cell decisions, then we can establish this hidden connection at least in principle via *maximum a posteriori* estimation of the intensity of cell decisions and the growth-related transformation.

---

<sup>1</sup>This is a biological time-varying curvilinear coordinate system that reflects shape and interior deformations of an organism induced by growth. Computational methods for its generation are proposed in [14].

We demonstrate inference results for the larval growth of the *Drosophila* wing disc using a sequence of micrographs of the Wingless expression pattern. The Wingless gene plays a primary role in patterning of adult fly structures, generating their dynamic geographic map [16]. There are also other key regulatory genes that promote morphogenesis, namely, Apterous and Vestigial [16]. We don't consider their expression patterns in our model in order to avoid additional complexity of the inference problem. Therefore, our macroscopic GRID characterization of the wing disc growth is based on a simplifying assumption that the dynamics of secondary key regulatory gene products does not depend on a growth process (cell divisions at larval stage).

We show that the estimated biological transformations attempt to register images of Wingless protein concentration in the initial and grown wing discs in a biologically meaningful way only for a certain order of magnitude of the prior energy coefficient. This weighting coefficient is a measure of independency between the observed dynamics of the gene expression pattern of interest and the growth process. Its values of order  $O(1)$  suggest that the growth process does not drive changes in the initial pattern of protein concentration into a registration with the observed pattern in the grown *Drosophila* wing disc. Evaluation of the prior energy coefficient follows from an experimental study of the nonlinear behavior of the cost function when treated as a function of the unknown weighting coefficient.

*Paper organization.* This paper is organized as follows. We introduce the GRID macroscopic growth integro-differential equation and the discrete macroscopic growth operator in Sections II and III. We show invariance of the inferred Poisson intensity parameter under the Euclidean transformations in Section IV. In Section V we study regularity properties of the macroscopic growth operator which are necessary for numerical stability of the inference. As a result, we impose a constraint on the intensity function that ensures its smoothness throughout the organism's domain.

In Section VI, following a methodology in computational anatomy of constructing a flow as a shortest path from one specimen to another, we formulate a *maximum a posteriori* estimation problem using the macroscopic growth law. Modeling the cost function as a Gibbs potential that is the sum of observation (or likelihood) and prior energies we describe a Gaussian observation model that relates the image data with the unknown intensity parameter. In Sections VII and VIII we consider prior models for the unobserved intensity function as a Gaussian random field and as a random field governing a spatial Poisson point process. We show that the latter choice is biologically motivated and consistent with the discretized version of the macroscopic growth law. For the latter prior model we study the performance of the direct estimation algorithm depending on the value of the prior energy coefficient in Section IX. Here we also find experimentally the value that yields 'best' inferences characterizing larval growth of the *Drosophila* wing disc, and in Section X we estimate the Poisson intensity field and the biological transformation of the growing wing disc from the sequence of micrographs of Wingless gene expression patterns. Appendix A explains major steps of the image inference algorithm. We conclude our paper with the summary in Section XI.

## II. GRID macroscopic growth law

Motivation for the GRID macroscopic growth law comes from the developmental biology of multicellular organisms. The phenomenon of biological growth manifests itself in a time-varying deformation of shape and internal structure of a developing organism. A multitude of elementary biological events such as cell divisions and/or enlargements, cell deaths and cell movements results in visible shape and interior changes of the growing multicellular organism and its suborganisms seen in images collected over time.

It seems natural to represent an underlying biological transformation on a large time scale by a diffeomorphic flow that evolves in time as a collective effect of a large number of cell decisions. The diffeomorphic property of the transformation preserves integrity of the deforming Darcyan curvilinear grid representing the growing multicellular structure as it does not allow local overlaps. To generate such flows we use the GRID macroscopic growth law introduced earlier [11], that is, the deterministic integro-differential equation emphasizing dependency of the solution on such GRID variables as the Poisson intensity of cell decisions and the relative rate of expansion/contraction. We keep the notation for GRID growth variables as consistent as possible with the seminal paper by U. Grenander “On the Mathematics of Growth” [11].

For an intuitive understanding of this equation we give a synthetic example of a macroscopic growth of an organism in the form of a unit disc. With preassigned space-dependent Poisson intensity of cell divisions and constant relative rate of expansion we run a few realizations of the macroscopic growth equation on the Darcyan grid of the disc. Initially, the Darcyan grid is the polar grid that deforms as the disc grows. Here, the Poisson intensity is the probability mass function assigned to all the nodes of the Darcyan coordinate grid. The center of mass is located at the origin of the coordinate system and the probability mass decays radially from it. As a result, organism’s central part grows faster than the rest of it (see Fig. 1).

When the GRID variables are known this equation predicts a typical growth pattern in the average sense. At each instant of time the growth increment is computed by averaging over all seed contributions. According to the isotropic GRID model the displacement field due to a single seed contribution,  $x(\xi_{seed})$ , is modeled as an exponential decay of the radial distance from the seed.

$$\begin{aligned} \Delta^{\xi_{seed}} x(\xi, t) &= k(x(\xi, t)) \cdot (x(\xi, t) - x(\xi_{seed}, t)) \cdot \exp\left(-\frac{\|x(\xi, t) - x(\xi_{seed}, t)\|^2}{step(x(\xi_{seed}, t))^2}\right) \quad (\text{II.1}) \\ &= k(x(\xi_{seed}, t)) \cdot \theta(x(\xi, t) - x(\xi_{seed}, t)). \end{aligned}$$

Here,

- (1)  $x(\xi, t) = (x_1(\xi, t), x_2(\xi, t))$  where  $\xi = (\xi_1, \xi_2) \in \Xi$  is a curvilinear Darcyan coordinate system of the organism’s domain, the  $X$ -field  $X(t) \in R^2$ .
- (2)  $\Xi$  is the Darcyan space of biological coordinates of cell decisions. Given the Darcyan grid with  $n$  radial and  $m$  angular coordinate curves the  $\Xi$ -space is a product of two finite sets of integers  $\{1, 2, 3, \dots, n\}$  and  $\{1, 2, 3, \dots, m\}$ .
- (3)  $k(x(\xi_{seed}, t))$  is the relative rate of expansion/contraction independent of the angle or direction of growth/decay.  $0 < k(x(\xi, t)) < 1$  implies expansion of the

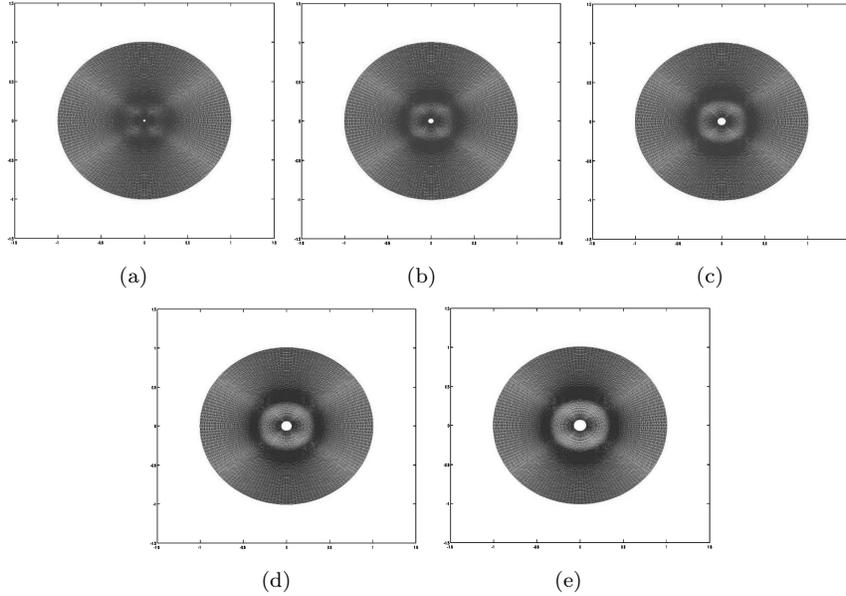


FIG. 1. (a-e) Example of macroscopic growth with space-dependent Poisson intensity. The Poisson intensity is the probability mass function which decreases with respect to radial distance. Result: the central part grows faster.

small area around  $x(\xi, t)$  and  $-1 < k(x(\xi, t)) < 0$  implies local contraction. So, small areas centered at seeds grow or decay isotropically.

- (4)  $step(x(\xi_{seed}, t))$  is the radius of influence of the seed. Since  $x(\xi, t)$  provides a computational grid with nonuniformly distributed nodes, the  $step$  parameter is set to be the Jacobian of the transformation  $x(t) = x(\xi, t)$ , i.e.,

$$step(x(\xi, t)) = \left| \frac{\partial(x_1, x_2)}{\partial(\xi_1, \xi_2)} \right| \text{ (see Fig. 2).}$$

In this way we compensate for the nonuniformity of grid quadrilaterals and establish independency of displacements in the sense that they don't overlap when they occur simultaneously in the neighborhoods of seeds located at the nodes of the Darcyan grid.

For a fixed seed  $x(\xi_{seed}, t)$  (II.1) defines a translation-invariant diffeomorphism  $\theta(x(\xi, t) - x(\xi_{seed}, t)) : R^2 \rightarrow R^2$ . At a time  $t$  the points  $x(\xi, t)$  of the  $X$ -field move relative to the location of that seed.

According to the stochastic version of the GRID model for growth on a fine time-scale, seeds are activated within the growing organism as an inhomogeneous space-time Poisson point process with the intensity parameter  $\lambda(\xi, t)$ . To approximate observed growth in instantaneous time we use the Law of Large Numbers. We let a number of cell decisions (seeds) that occur randomly and independently in the growing organism's domain become arbitrarily large at each instant of time. Then the growth-induced velocity is the average value of the velocity field taken over all seeds. Thus, we formulate the macroscopic

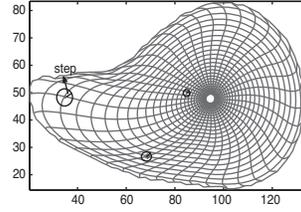


FIG. 2. Illustration of the step parameter as a function defined on the computational Darcyan grid of the *Drosophila* wing disc.

growth law in the integro-differential form

$$\begin{aligned} \frac{\partial x(\xi, t)}{\partial t} &= \int_{\xi_{seed} \in \Xi} \Delta^{\xi_{seed}} x(\xi, t) F(d\xi_{seed}) \\ &= \int_{\xi_{seed} \in \Xi} k(x(\xi_{seed}, t)) \cdot \theta(x(\xi, t) - x(\xi_{seed}, t)) \cdot f(x(\xi_{seed}, t)) d\xi_{seed}, \end{aligned} \quad (\text{II.2})$$

with initial conditions  $x(\xi, 0) = x(\xi)$  and the Poisson measure of cell density in the  $X$ -field  $f(x(\xi_{seed}, t))$  evaluated at seed locations.

It is important to realize that  $f(x(\xi_{seed}, t))$  is expressed in absolute space coordinates  $x$  that depend on Darcyan coordinates  $\xi_{seed}$  of local gene activity sites. Then by the law of probability conservation,

$$f(x) = f(x(\xi, t)) = \frac{\lambda(\xi, t)}{J(x(\xi, t))}, \quad (\text{II.3})$$

where  $J(x(\xi, t))$  is the Jacobian of the transformation  $x = x(\xi)$  at a time  $t$ . The Poisson intensity measure expressed in Darcyan coordinates  $\xi_{seed} \in \Xi$  is an intrinsic growth property of an organism represented by  $\lambda(\xi_{seed})$ . It influences an organism's overall shape changes causing local contractions or expansions of its cellular field viewed as a continuum  $X$ -field.

To characterize macroscopic growth, U. Grenander introduced a bounded function  $a(\xi_{seed}, t)$  called the growth magnitude [11]

$$a(\xi_{seed}, t) = k(x(\xi_{seed}, t)) \cdot f(x(\xi_{seed}, t)) = k(x(\xi_{seed}, t)) \frac{\lambda(\xi_{seed}, t)}{J(x(\xi_{seed}, t))}. \quad (\text{II.4})$$

It simplifies the statement of the macroscopic growth law but hides the important growth parameter  $\lambda(\xi_{seed}, t)$  within itself. We would like to establish the dependency of the flow  $x(\xi, t)$  on the intrinsic intensity of cell decisions.

Since there are seemingly many biological events contributing to growth it is reasonable to approximate seed deformation effects by the local rate of expansion/contraction  $k(x(\xi_{seed}, t))$  that is constant in magnitude and varying in sign throughout the  $X$ -field. That is,

$$k(x(\xi_{seed}, t)) = \text{sgn}(k) \cdot \text{const.}$$

Let  $k(x(\xi_{seed}, t)) \in \{-1, 1\}$  and let  $\lambda$  absorb this constant for simplicity. Then the sign of  $\lambda(\xi_{seed}, t)$  will define the nature of the seed  $\xi_{seed}$  (growth or decay). The Poisson intensity

of events in the Darcyan space is  $|\lambda(\xi_{seed}, t)|$  and the growth magnitude becomes

$$a(\xi_{seed}, t) = \frac{\lambda(\xi_{seed}, t)}{J(x(\xi_{seed}, t))}. \quad (\text{II.5})$$

We reformulate the macroscopic growth equation with respect to the  $\lambda$ -field

$$\frac{\partial x(\xi, t)}{\partial t} = \int_{\xi_{seed} \in \Xi} \theta(x(\xi, t) - x(\xi_{seed}, t)) \cdot \frac{\lambda(\xi_{seed}, t)}{J(x(\xi_{seed}, t))} d\xi_{seed}. \quad (\text{II.6})$$

The macroscopic growth equation (II.6) can be rewritten in the operator form as follows:

$$\begin{aligned} \frac{\partial x(\xi, t)}{\partial t} &= (\Theta\lambda)(x(\xi, t)) \\ &= \int_{\xi_{seed} \in \Xi} \frac{\theta(x(\xi, t) - x(\xi_{seed}, t))}{J(x(\xi_{seed}, t))} \cdot \lambda(\xi_{seed}, t) d\xi_{seed}, \end{aligned} \quad (\text{II.7})$$

where

$$(\Theta\lambda)(x(\xi, t)) = \begin{pmatrix} \Theta_{x_1} \lambda \\ \Theta_{x_2} \lambda \end{pmatrix}. \quad (\text{II.8})$$

$\Theta$  is an integral operator acting on a scalar  $\lambda$ -field,  $\lambda : (\xi_1, \xi_2) \in \Xi \rightarrow R$  that assigns a vector to each point of the  $\lambda$ -field, that is,  $(\Theta\lambda) : R \rightarrow R^2$ . Thus, the  $\Theta$  operator returns the average velocity field at a time  $t$ .

### III. The discrete macroscopic growth operator

In practical applications we implement a discrete version of the macroscopic growth operator  $\Theta$ . Let the Darcyan coordinate system of an organism be a curvilinear grid with  $n$  radial and  $m$  angular coordinate curves. Then the organism's domain is represented by a configuration of the Darcyan grid nodes

$$\{x(\xi_1, \xi_2, t) : \xi_1 \in \{1, 2, \dots, n\}, \xi_2 \in \{1, 2, \dots, m\}\}.$$

For simplicity of notation the Darcyan coordinate system is referred to as  $\{x(\xi_i, t), 1 \leq \xi_i \leq N, N = m \cdot n\}$ .

We consider the space-time discretization of the macroscopic growth equation. At a time  $t = j$  a finite-dimensional approximation  $\Delta x(\xi, l)$  of the infinite-dimensional displacement field is

$$\Delta x(\xi, j) = (\Theta_N \lambda) = \sum_{\xi_{seed} \in \Xi} \frac{\theta(x(\xi, (j-1)) - x(\xi_{seed}, (j-1)))}{J(x(\xi_{seed}, (j-1)))} \cdot \lambda(\xi_{seed}, j), \quad (\text{III.1})$$

where  $\xi, \xi_{seed} \in \Xi = \{1, 2, \dots, N\}$  and

$$(\Theta_N \lambda)(x(\xi, j)) = \begin{pmatrix} \Theta_{N_{x_1}} \lambda \\ \Theta_{N_{x_2}} \lambda \end{pmatrix}. \quad (\text{III.2})$$

As  $N \rightarrow \infty$ ,  $\Theta_N \rightarrow \Theta$ .  $\Theta_N$  is a discrete macroscopic growth operator,

$$\begin{aligned} \Theta_N \lambda(x(\xi_i, j)) &= \sum_{l=1}^N (x(\xi_i, j-1) - x(\xi_l, j-1)) \\ &\cdot \exp\left(-\frac{\|(x(\xi_i, j-1) - x(\xi_l, j-1))\|^2}{step(x(\xi_l, j-1))^2}\right) \cdot \frac{\lambda(\xi_l, l)}{J(x(\xi_l, j-1))}. \end{aligned} \quad (\text{III.3})$$

For each grid node  $i$ ,  $\Theta_N \lambda(x(\xi_i, l)) = \Delta x(\xi_i, l)$ , where  $\Delta x(\xi_i, l) = x(\xi_i, l) - x(\xi_i, l - 1)$  is a 2D displacement vector. According to (III.2)-(III.3) the displacement field is computed by multiplying each component  $\Theta_{N_{x_k}}$  of a 2D array  $\Theta_N$  by an  $N$ -dimensional vector  $\lambda$ . Clearly,  $\Theta_{N_{x_k}}$  is an  $N \times N$  matrix whose  $i^{\text{th}}$  row entries represent displacements of a seed  $\xi_i$  resulting from the activation of seeds  $\xi_l$ ,  $1 \leq \xi_l \leq N$ .

DEFINITION III.1. The discrete macroscopic growth operator  $\Theta_N$  at a time  $t = l$  associated with the  $n \times m$  Darcyan grid of an organism is a two-dimensional array of

$N \times N$  matrices  $\begin{pmatrix} \Theta_{N_{x_1}} \\ \Theta_{N_{x_2}} \end{pmatrix}$  with  $N = n \cdot m$  and entries

$$\Theta_{N_{x_k} i_j} = \frac{1}{J(x_k(\xi_j, l - 1))} \cdot (x_k(\xi_i, l - 1) - x_k(\xi_j, l - 1)) \exp\left(-\frac{\|x_k(\xi_i, l - 1) - x_k(\xi_j, l - 1)\|^2}{\text{step}(x_k(\xi_j, l - 1))^2}\right). \quad (\text{III.4})$$

#### IV. Invariance of the inference of the intensity parameter under the Euclidean group of transformations

In biology, measurements of growth parameters of a developing organism do not depend on whether an observer rotates or translates its coordinate system in space. Therefore, it is important that the  $\lambda$ -field inferred from images using the macroscopic growth equation is independent of the rigid motion of an organism. In this section we show that the translation-invariant model of the deformation field due to a single seed given in (II.1) leads to the invariance of the inference of  $\lambda(\xi, t)$  with respect to spatial translations, rotations and reflections.

PROPOSITION IV.1. The inference of  $\lambda(\xi, t)$  is invariant under the Euclidean group of transformations  $E(2)$ .

*Proof.* We prove that the inferred  $\lambda$ -field is invariant under the spatial translations, rotations and reflections of the Darcyan coordinate system of an organism.

*Case 1: Translation.* We apply a spatial transformation to the Darcyan coordinate system of an organism  $x(\xi, t) \rightarrow x(\xi, t) + b$ . For a fixed  $\xi$ ,  $x(\xi)$  is a 2D vector with  $x_1$  and  $x_2$  absolute coordinates. Let  $\tilde{x}(\xi, t) = x(\xi, t) + b$ . Clearly, the Jacobian of the transformation  $\tilde{x}(\xi, t)$  with respect to  $\xi$  does not change,  $J(\tilde{x}(\xi, t)) = J(x(\xi_{seed}, t))$ , as well as the *step* parameter since it is equal to the Jacobian.

Also, the displacement field for a fixed seed  $\xi_{seed} \in \Xi$  remains the same due to its translation-invariance property

$$\begin{aligned} \theta(\tilde{x}(\xi, t) - \tilde{x}(\xi_{seed}, t)) &= (\tilde{x}(\xi, t) - \tilde{x}(\xi_{seed}, t)) \\ &\cdot \exp\left(\frac{-\|\tilde{x}(\xi, t) - \tilde{x}(\xi_{seed}, t)\|^2}{\text{step}(\tilde{x}(\xi_{seed}, t))^2}\right) \\ &= \theta(x(\xi, t) - x(\xi_{seed}, t)). \end{aligned}$$

Therefore, at a time  $t = j$  we obtain the average displacement field

$$\begin{aligned}\Delta\tilde{x}(\xi, j) &= (\Theta_N\lambda)(\tilde{x}(\xi, j)) \\ &= \sum_{\xi_{seed} \in \Xi} \theta(\tilde{x}(\xi, j-1) - \tilde{x}(\xi_{seed}, j-1)) \cdot \frac{\lambda(\xi_{seed}, j)}{J(\tilde{x}(\xi_{seed}, j-1))} \\ &= \sum_{\xi_{seed} \in \Xi} \theta(x(\xi, j-1) - x(\xi_{seed}, j-1)) \cdot \frac{\lambda(\xi_{seed}, j-1)}{J(x(\xi_{seed}, j-1))} = \Delta x(\xi, j).\end{aligned}$$

Then  $\lambda(\xi, j) = \Theta_N^{-1}\Delta\tilde{x}(\xi, j) = \Theta_N^{-1}\Delta x(\xi, j)$ .

*Case 2: Rotation.* For the rotation transformation  $x(\xi, t) \rightarrow Rx(\xi, t)$ , where  $R = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$ , we let  $\tilde{x}(\xi, t) = Rx(\xi, t)$ . Then on the LHS of (III.1) at a time  $t = j$  we have

$$\Delta\tilde{x}(\xi, j) = \Delta Rx(\xi, j) = R\Delta x(\xi, j). \quad (\text{IV.1})$$

Since rotation of the coordinate system preserves Euclidean distances  $\|(\tilde{x}(\xi, j-1) - \tilde{x}(\xi_{seed}, j-1))\| = \|(x(\xi, j-1) - x(\xi_{seed}, j-1))\|$  between seeds and the areas around them,  $J(\tilde{x}(\xi_{seed}, j-1)) = J(x(\xi_{seed}, j-1)) = \text{step}(x(\xi_{seed}, j-1))$ , then on the RHS of the discrete macroscopic growth equation we obtain

$$\begin{aligned} &(\Theta_N\lambda)(Rx(\xi, j)) \\ &= \sum_{\xi_{seed} \in \Xi} (\tilde{x}(\xi, j-1) - \tilde{x}(\xi_{seed}, j-1)) \cdot \exp\left(\frac{-\|(x(\xi, j-1) - x(\xi_{seed}, j-1))\|^2}{\text{step}(x(\xi_{seed}, j-1))^2}\right) \\ &\quad \cdot \frac{\lambda(\xi_{seed}, j)}{J(x(\xi_{seed}, j-1))} \\ &= R \left( \sum_{\xi_{seed} \in \Xi} (x(\xi, j-1) - x(\xi_{seed}, j-1)) \cdot \exp\left(\frac{-\|(x(\xi, j-1) - x(\xi_{seed}, j-1))\|^2}{\text{step}(x(\xi_{seed}, j-1))^2}\right) \right) \\ &\quad \cdot \frac{\lambda(\xi_{seed}, j)}{J(x(\xi_{seed}, j-1))} \\ &= R((\Theta_N\lambda)(x(\xi, j))). \end{aligned}$$

Equating the LHS (IV.1) and the RHS leads to

$$\Delta\tilde{x}(\xi, j) = (\Theta_N\lambda)(\tilde{x}(\xi, j)) = R((\Theta_N\lambda)(x(\xi, j))) = R\Delta x(\xi, j). \quad (\text{IV.2})$$

Equation (IV.2) implies the solution to the inverse problem

$$\lambda(\xi, j) = \Theta_N^{-1}\Delta\tilde{x}(\xi, j) = \Theta_N^{-1}R^{-1}R\Delta x(\xi, j) = \Theta_N^{-1}x(\xi, j).$$

*Case 3: Reflection.* By the same argument as in case 2 the  $\lambda$ -field remains unchanged under the reflection transformation  $x(\xi, t) \rightarrow Rx(\xi, t)$ , where  $R = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$ .  $\square$

The proposition reflects the fact that in this model a developing organism does not know the absolute space.

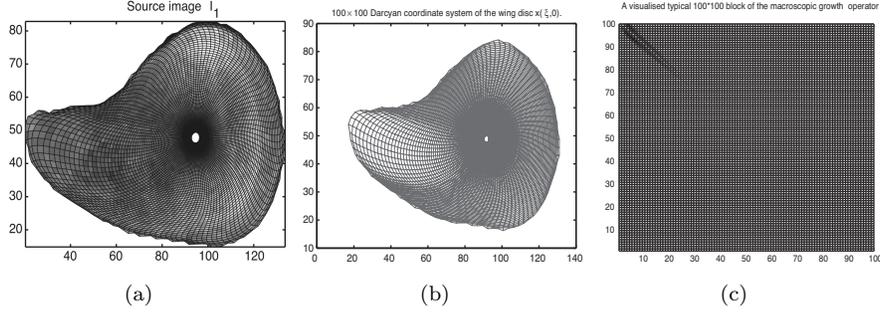


FIG. 3. (a) Wingless gene expression pattern in the *Drosophila* wing disc interpolated onto (b) a  $100 \times 100$  Darcyan coordinate system  $x(\xi, 0)$ , (c) Typical  $100 \times 100$  'tile' of the macroscopic growth matrix  $\Theta_{10,000_{x_1}}$ .

## V. Regularity properties of the macroscopic growth operator

We study regularity properties of  $\Theta_N$  to address a question of the numerical stability of the solution to the inverse problem  $\lambda(\xi, t) = \Theta_{N_{x_k}}(\Delta x(\xi, t))^{-1}$ , where  $k = 1$  or  $2$ . Using (III.4) we compute the macroscopic growth matrix  $\Theta_{10,000}$  associated with a  $100 \times 100$  Darcyan grid of the *Drosophila* wing disc  $x(\xi)$  (see Figs. 3(a)-3(b)). Figure 3(c) provides an insight into the structure of the macroscopic growth matrix  $\Theta_{10,000}$ .

In Figure 3(c) negative and positive displacements are the entries above and below the main diagonal of the matrix, respectively. Their non-zero values are shown in black. The structure of the macroscopic growth matrix is reminiscent of a skew-symmetric matrix whose elements satisfy  $\Theta_{ij} = -\Theta_{ji}$ . Although  $\Theta_{N_{x_k}}$  is not skew-symmetric, as dictated by its definition (III.4), we can argue that for the seeds (nodes of the Darcyan grid)  $\xi_i$  and  $\xi_j$  located close to one another,

$$\begin{aligned} J(x_k(\xi_i, l-1)) &\approx J(x_k(\xi_j, l-1)), \\ \text{step}(x_k(\xi_i, l-1)) &\approx \text{step}(x_k(\xi_j, l-1)), \\ \Theta_{N_{x_k}}{}_{ij} &\approx (x_k(\xi_i, l-1) - x_k(\xi_j, l-1)) \cdot \text{const}. \end{aligned}$$

Since the displacement decays exponentially to zero as the distance between  $\xi_i$  and  $\xi_j$  increases we have

$$\Theta_{N_{x_k}}{}_{ij} \approx 0.$$

Therefore, to a certain order of accuracy, the macroscopic growth matrix is approximately skew-symmetric. Then an eigenvalue decomposition can be applied to such an approximation since it is a normal matrix. That is,

$$\Theta_{N_{x_k}} \approx \tilde{\Theta}_{N_{x_k}} \text{ and } \tilde{\Theta}_{N_{x_k}} = U\Lambda U^T,$$

where  $\Lambda$  is a block-diagonal matrix of the eigenvalues of  $\tilde{\Theta}_{N_{x_k}}$  and  $U$  is an orthogonal matrix whose columns are the eigenvectors of  $\tilde{\Theta}_{N_{x_k}}$ .

In order that the numerical solution to the inverse problem be stable,  $\Theta_{N_{x_k}}$  should be well-conditioned. A well-conditioned matrix ensures that  $\lambda(\xi, t)$  does not grow without

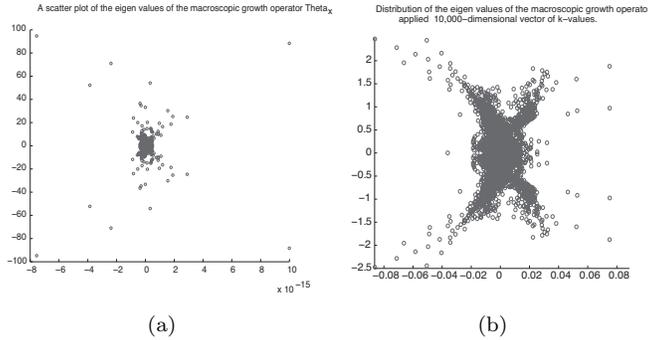


FIG. 4. (a) Scatter plot of the eigenvalues of the macroscopic growth operator  $\Theta_{10,000,x_1}$  and (b) a close-up view of the eigenvalue distribution in the complex plane.

bounds in response to small changes in  $\Delta x(\xi, t)$ . We show below that  $\Theta_{10,000,x_k}$  associated with the Darcyan grid of the *Drosophila* wing disc (see Fig. 3) is ill-conditioned.

Eigenvalues of the  $\Theta_{10,000,x_1}$  matrix come in conjugate pairs of complex numbers (see Fig. 4) similar to conjugate pairs of purely imaginary eigenvalues of the skew-symmetric matrix. The scatter plots in Figure 4 show that the eigenvalues  $\mu_i$ ,  $1 \leq i \leq 10,000$ , accumulate near zero with only a few of them having a large magnitude. Since  $\Theta_{10,000,x_1}$  is approximately the skew-symmetric matrix  $\tilde{\Theta}_{10,000,x_1}$  we can evaluate the condition number of  $\Theta_{10,000,x_1}$  using a formula for the condition number of a normal matrix

$$\kappa(\tilde{\Theta}) = \left| \frac{\mu_{max}(\tilde{\Theta})}{\mu_{min}(\tilde{\Theta})} \right|.$$

(The condition number measures the rate of change in  $\lambda(\xi, t)$  with respect to a change in  $\Delta x(\xi, t)$ .) Since the smallest eigenvalue  $\mu_{min} = 0$  and  $\mu_{max} > 0$ ,  $\kappa(\Theta_{10,000,x_1}) = \infty$  and we conclude that the solution to the linear system of equations

$$\Delta x(\xi, t) = \Theta_{10,000,x_1} \lambda(x(\xi, t)), \quad \xi \in \Xi = \{1, 2, \dots, 10,000\}$$

is numerically unstable.

In order to obtain a stable solution we must impose additional assumptions on the unknown  $\lambda$ -field. Statistically, we might assume, *a priori*, that  $\lambda(\xi, t) \in R^N \times [0, T]$  is a random variable with a multivariate normal probability distribution at any time  $t$  (a random field). It is reasonable to take the mean  $\langle \lambda(t) \rangle = \frac{1}{N}$  reflecting the fact that all cells have an equal chance to divide. For simplicity, we assume that each component  $\lambda(\xi_i)$  is independent with standard deviation  $\sigma_\lambda$ .

Since the image data are subject to errors, we can assume that the difference between the inferred image of a grown organism (target) and the target image,  $I_1(x^{-1}(\xi, t)) - I_2(x(\xi, 0))$ , is an error that is an  $N$ -dimensional normally distributed random vector with components  $e_i \sim \mathcal{N}(0, \sigma_I^2)$ . Using Bayes's theorem we can find the most likely solution given the image data and the *a priori* distribution of  $\lambda(\xi, t)$  that is equivalent to a Tikhonov-regularized solution. This is the subject of Section VII.

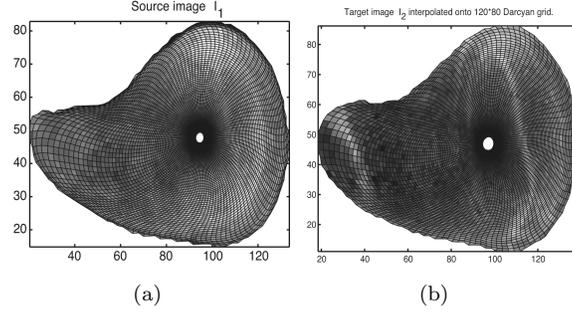


FIG. 5. (a) Source  $I_1(x(\xi))$  and (b) Target  $I_2(x(\xi), T)$  ( $\approx 11$  hours later) images of the Wingless expression pattern in the *Drosophila* wing disc interpolated onto the Darcyan grid of the wing disc  $x(\xi)$  in  $I_1$ .

Using the Bayesian view of the unknown  $\lambda(\xi, t)$  we might think of its absolute value as a random field controlling occurrences of cell decisions in a developing organism. In order to connect to the true biological process of growth we should not prescribe a probability distribution thus weakening the constraints on the  $\lambda$ -field. We propose representing an *a priori* knowledge about the Poisson intensity parameter  $|\lambda(\xi, t)|$  as a sample-density function for a spatial Poisson point process in Section VIII. Since such an *a priori* model does not lead to a smooth solution we can regularize it by introducing a penalty function in the form of a squared  $L^2$ -norm of the operator  $\Phi = \sum_i a_i \nabla^{2 \cdot q_i}$  known as *Good's roughness* [17],

$$\phi(\lambda) = \|\Phi\lambda\|_2^2.$$

Taking only  $q_1 = 1/2$  we obtain the total variation integral

$$\phi(\lambda) = \|\nabla\lambda\|_2^2 = \int_{\xi \in \Xi} |\nabla\lambda(\xi, t)|^2 d\xi.$$

Since  $\xi$  spans a finite range of integers we use the total variation integral discretized in  $\Xi$ -space to constrain a *maximum a posteriori* estimate of  $\lambda(\xi, t)$  to be a smooth function in the organism's domain.

## VI. Inference problem formulation using the macroscopic growth law

Given a pair of images  $I_1(x(\xi, 0))$ (source) and  $I_2(x(\xi), T)$ (target) of an organism growing over a finite time period  $[0, T]$  (see Fig. 5) we would like to estimate the GRID growth parameter  $\lambda(\xi, t)$  and the underlying biological transformation  $x(\xi, t)$ ,  $0 \leq t \leq T$ . We follow a general approach of constructing the growth-induced transformation as a diffeomorphic flow with the least energy property adopted in computational anatomy. Generated by the macroscopic growth equation, such a transformation depends on the optimal value of  $\lambda(\xi, t)$ .

Using Bayes's theorem the optimal *maximum a posteriori* (MAP) estimate of  $\lambda(\xi, t)$  is the minimum of the following posterior energy (Gibbs potential),

$$E_{post} = E_{likelihood}(\lambda(\xi, t)) + E_{prior}(\lambda(\xi, t)),$$

resulting in the following minimization problem,

$$\lambda(\hat{\xi}, t) = \arg \min_{\lambda \in R^N \times [0, T]} \{E_{likelihood}(\lambda(\xi, t)) + E_{prior}(\lambda(\xi, t))\}, \quad (VI.1)$$

where  $E_{likelihood}(\lambda(\xi, t)) = -\ln(p(I_2(x(\xi, t)) | \lambda(\xi, t)))$  and  $E_{prior} = -\ln(p(\lambda(\xi, t)))$ . Here,  $p(I_2(x(\xi, t)) | \lambda(\xi, t))$  is the likelihood function that relates observed pixel values with the growth parameter  $\lambda(\xi, t)$  and  $p(\lambda(\xi, t))$  is the prior probability density function (Gibbs measure).

At a time  $t$ ,  $x(\xi, t)$  is constrained to be the solution to the discretized macroscopic growth equation,

$$\begin{aligned} x(\xi, t) - x(\xi, 0) &= \sum_{j=1}^t \Delta x(\xi, j) \\ &= \sum_{j=1}^t \sum_{\xi_{seed} \in \Xi} \frac{\theta(x(\xi, (j-1)) - x(\xi_{seed}, (j-1)))}{J(x(\xi_{seed}, (j-1)))} \cdot \lambda(\xi_{seed}, j) \end{aligned} \quad (VI.2)$$

and  $\lambda(\xi, t)$  is constrained to satisfy  $\sum_{\xi \in \Xi} |\nabla \lambda(\xi, t)|^2 = 0$  if increasing smoothness is needed.

As time  $t \rightarrow T$  the posterior energy  $E_{post}$  will decrease toward a local minimum whose value is not necessarily zero.

We view the target image  $I_2$  as a result of the evolution of the source image  $I_1$  over the time period of growth  $[0, T]$ , that is,

$$I_2 = I_2(x(\xi), T) = I_1(x(\xi, T)^{-1}) + \epsilon, \quad (VI.3)$$

where  $x(\xi, T)$  is the solution to the macroscopic growth equation at time  $T$  and  $\epsilon \sim N(0, \sigma_I^2)$  is independent, additive Gaussian noise that models the observation  $I_2$  as a degraded version of the estimated image. As such, the target image  $I_2$  becomes a sample from the Gaussian process with the mean being the source image  $I_1$  under the time-dependent deformation. The probability distribution of  $I_2(x(\xi, t)) = \{I_{2_1}, I_{2_2}, \dots, I_{2_N}\}$  conditional on  $\lambda(\xi, t)$  is

$$\begin{aligned} p(I_2(x(\xi, t)) | \lambda(\xi, t)) &= \frac{J(x(\xi))}{(\sqrt{2\pi\sigma_I^2})^N} \\ &\cdot \exp\left(-\frac{1}{2\sigma_I^2} \cdot \sum_{\xi \in \Xi} (I_2(x(\xi), T) - I_1(x(\xi, T)^{-1}))^2\right). \end{aligned} \quad (VI.4)$$

Thus, we are led to the discrete optimal control problem of the form,

$$\hat{\lambda}(\xi, T) = \arg \min_{\lambda} \left[ \sum_{\xi \in \Xi} \frac{1}{2\sigma_I^2} (I_2(x(\xi), T) - I_1(x(\xi, T)^{-1}))^2 \cdot J(x(\xi)) + E_{prior}(\lambda(\xi, T)) \right] \quad (VI.5)$$

subject to

$$x(\xi, T) - x(\xi, 0) = \sum_{j=1}^T \sum_{\xi_{seed} \in \Xi} \frac{\theta(x(\xi, (j-1)) - x(\xi_{seed}, (j-1)))}{J(x(\xi_{seed}, (j-1)))} \cdot \lambda(\xi_{seed}, j) \quad (VI.6)$$

with the initial conditions  $x(\xi, 0) = x(\xi)$  and (if needed)

$$\sum_{\xi \in \Xi} |\nabla \lambda(\xi, T)|^2 = 0. \quad (\text{VI.7})$$

Here,  $\{\lambda(\xi, t)\}_{t=1}^T$  is a sequence of optimal controls. We seek an estimate of the  $\lambda$ -field,  $\hat{\lambda}(\xi, T) = \lim_{t \rightarrow T} \lambda(\xi, t)$ , such that  $I_1((x(\xi), T)^{-1}) \sim I_2(x(\xi), T)$ .

The additional equality constraint (VI.7) could be removed by means of penalty functions [17]. Then the optimization problem (VI.5) is formulated in the unconstrained form

$$\begin{aligned} \hat{\lambda}(\xi, T) = \arg \min_{\lambda \in R^N \times [0, T]} & \sum_{\xi \in \Xi} \frac{1}{2\sigma_I^2} (I_2(x(\xi), T) - I_1(x(\xi, T)^{-1}))^2 \cdot J(x(\xi)) \\ & + E_{prior}(\lambda(\xi, T)) + \sum_{\xi \in \Xi} |\nabla \lambda(\xi, T)|^2. \end{aligned} \quad (\text{VI.8})$$

For the penalty function we have to compute the gradient of the scalar  $\lambda$ -field at each node of the Darcyan grid  $\xi = (\xi_1, \xi_2)$ ,  $1 \leq \xi_1 \leq n$ ,  $1 \leq \xi_2 \leq m$ ,  $N = n \times m$ ,

$$\nabla \lambda(\xi, T) = \frac{\partial \lambda(\xi, t)}{\partial \xi_1} + \frac{\partial \lambda(\xi, t)}{\partial \xi_2}. \quad (\text{VI.9})$$

Given a model of the cost function as the Gibbs potential that is the sum of the likelihood and prior energies, the challenge is to find an appropriate prior knowledge representation of the growth parameter  $\lambda(\xi, t)$ .

## VII. The GRID growth parameter $\lambda(\xi, t)$ as a Gaussian random field

In the Bayesian view, the source of the growth-induced deformation is an unobserved space-time stochastic process  $\{\lambda(\xi, t) : \xi \in \Xi, 0 \leq t \leq T\}$  defined on the planar lattice  $\Xi$ , called a *random field*. At a particular instant in time we can consider such a process as a purely spatial process reflecting the fact that there is a large number of cell decisions (equal to the dimension of the problem  $N$ ) occurring in the growing organism's domain according to the macroscopic growth law.

Assuming that on average cell decisions are equally likely to happen everywhere in the organism's domain we can define the mean intensity  $\lambda$ -field as  $\langle \lambda(\xi, t) \rangle = \frac{1}{N}$  for  $\forall t \in [0, T]$ . (This assumption is actually true for the larval development of the *Drosophila* wing disc.) For simplicity we let all elements  $\lambda(\xi_i, t)$ ,  $i \in 1, 2, \dots, N$  be stochastically independent. Then we can model each  $\lambda(\xi_i, t)$  as a Gaussian distributed random variable with standard deviation  $\sigma_\lambda$ , that is,  $\lambda(\xi_i, t) \sim \mathcal{N}(\frac{1}{N}, \sigma_\lambda^2)$ . So, at a time  $t$ ,  $\lambda(\xi, t) = \Lambda(\xi)$  is a realization of the Gaussian random field with the probability measure

$$p(\Lambda(\xi)) = \frac{1}{(\sqrt{2\pi\sigma_\lambda^2})^N} \exp\left\{-\frac{1}{2\sigma_\lambda^2} \cdot \sum_{\xi \in \Xi} (\Lambda(\xi) - \frac{1}{N})^2\right\}. \quad (\text{VIII.1})$$

From (VIII.1) we define the prior energy

$$E_{prior}(\lambda(\xi, t)) = \frac{1}{2\sigma_\lambda^2} \cdot \sum_{\xi \in \Xi} (\lambda(\xi, t) - \frac{1}{N})^2 \quad (\text{VIII.2})$$

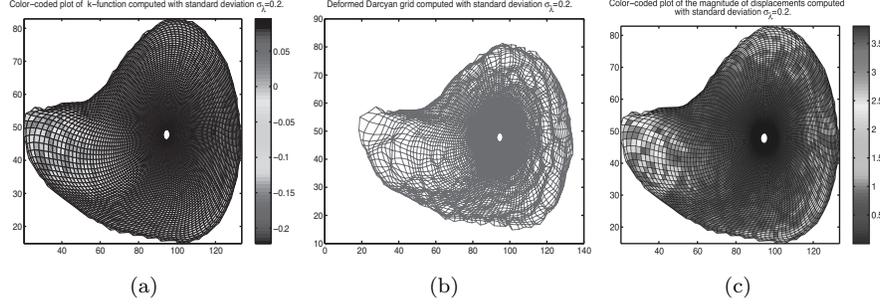


FIG. 6. (a) Color-coded plot of the optimal  $\lambda(\xi, T)$ -field, (b) Deformed Darcy grid of the wing disc given in  $I_1(x(\xi, 0))$ , and (c) Color-coded plot of the magnitude of the displacements for the image pair  $(I_1, I_2)$  (Fig. 5) and for  $\alpha = 0.5$ .

and formulate the resulting optimal control problem

$$\begin{aligned} \hat{\lambda}(\xi, T) = \arg \min_{\lambda \in R^N \times [0, T]} & \frac{1}{2\sigma_I^2} \cdot \sum_{\xi \in \Xi} (I_2(x(\xi), T) - I_1(x(\xi, T)^{-1}))^2 \cdot J(x(\xi)) \\ & + \frac{1}{2\sigma_\lambda^2} \cdot \sum_{\xi \in \Xi} (\lambda(\xi, t) - \frac{1}{N})^2, \end{aligned} \quad (\text{VIII.3})$$

where  $x(\xi, T)$  is subject to the macroscopic growth equation (VI.6). The optimization problem (VIII.3) is equivalent to

$$\begin{aligned} \hat{\lambda}(\xi, T) = \arg \min_{\lambda \in R^N \times [0, T]} & \sum_{\xi \in \Xi} (I_2(x(\xi), T) - I_1(x(\xi, T)^{-1}))^2 \cdot J(x(\xi)) \\ & + \|\Gamma[\lambda(\xi, t) - \frac{1}{N}]\|^2, \end{aligned} \quad (\text{VIII.4})$$

where  $\Gamma = \alpha I$  is a Tikhonov matrix with the Tikhonov factor  $\alpha = \frac{\sigma_I}{\sigma_\lambda}$ . Thus,  $\lambda(\xi, t)$  is a Tikhonov-regularized solution.

Observe that the larger the Tikhonov factor is, the smaller the variation of  $\lambda(\xi, t)$  is throughout the organism's domain. With a choice of  $120 \times 80$  Darcy grid of the *Drosophila* wing disc seen in  $I_1(x(\xi))$  (see Fig. 5(a)), an initial guess on the  $\lambda$ -field  $\lambda(\xi, 0) = \frac{1}{N}$ ,  $N = 120 \times 80$ , and  $\alpha = 0.5$ , we have implemented the direct estimation algorithm based on the Polak-Ribière conjugate gradient method (for computational details, see Appendix A) to solve the optimization problem (VIII.4). In Figure 6 are shown the estimates of the scalar  $\lambda$ -field, growth-induced transformation and the magnitude of the displacements obtained after 30 iterations.

REMARK VIII.1. Comparing the source and target images  $I_1$  and  $I_2$  given in Figure 5, we observe that a grey level mismatch is mostly present near the edge of the anterior part and throughout the dorsal part of the wing disc. The density of the Darcy grid quadrilaterals was increased by adding more angular grid curves and decreasing the average distance between evolving level sets (radial closed grid curves) closer to the boundary of the wing disc. Since there is no difference in the grey level content of both

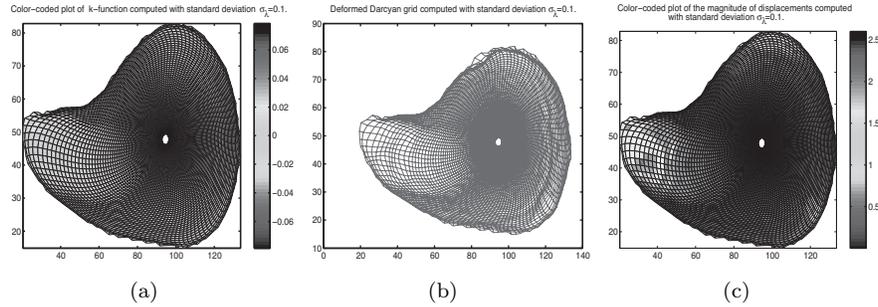


FIG. 7. (a) Color-coded plot of the optimal  $\lambda(\xi, T)$ -field, (b) Deformed Darcy grid of the wing disc given in  $I_1(x(\xi, 0))$ , and (c) Color-coded plot of the magnitude of the displacements for the image pair  $(I_1, I_2)$  (Fig. 5) and for  $\alpha = 1$ .

images in the neighborhood of the origin, propagation of level sets into this area was halted.

As expected,  $\lambda(\xi, T)$  is oscillating wildly throughout the domain causing local expansions of the initial Darcy grid combined with local contractions as seen from Figures 6(a) and 6(b). The magnitude of the oscillations is small, and the biological transformation  $x(\xi, t)$  estimated from  $\lambda(\xi, t)$ , where  $0 \leq t \leq T$ , preserves smoothness of the Darcy grid curves as seen from Figure 6(b). Such an estimate can describe biological growth with small regions in dark gray designating cell division locations. It is possible that the dividing cells push out the neighboring cells causing local contractions of the cellular field (shown in light gray). Regions of faster growth shown in dark gray appear like isolated cell division sites distributed uniformly throughout the wing disc. This is the consequence of the stochastic independency and identical distribution of elements of the random field  $\lambda(\xi, t)$ .

The magnitudes of the displacements appear in a pattern of concentric circles that are spread out in the interior domain excluding the central part of the wing disc where there are no changes in Wingless protein concentration. The maximal magnitudes shown in red are achieved at the tips of anterior and dorsal parts of the wing disc (see Fig. 6(c)).

To obtain a smooth behavior of the optimal  $\lambda$ -field we increase the Tikhonov factor to 1. Minimization of the posterior energy (VIII.4) with  $\alpha = 1$  leads to the desired smoothness of  $\hat{\lambda}(\xi, T)$  resulting in a lesser overall deformation effect as seen from Figures 7(a) and 7(b). The regions of high cell activity (cell divisions, cell shape change) appear restricted to the tip of the anterior part and the dorsal part of the wing disc. These are subregions of the wing disc where the highest Wingless protein concentration level mismatch is observed (see Fig. 5) and where the estimated magnitudes of displacements are the largest (see Fig. 7(c)). On the other hand, increasing the smoothness properties of  $\lambda(\xi, t)$  leads to a less satisfactory estimate of the target image  $I_2$  (see Fig. 8(b)). An image estimate shown in Figure 8(a) is a better result in terms of contrast present in the dorsal part of the wing disc (see Fig. 8(c)) and full reconstruction of the area near the tip of the anterior part. In both cases the estimated biological transformation tends to blur

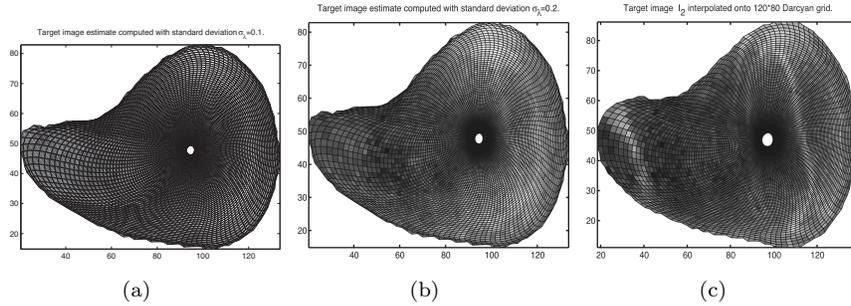


FIG. 8. Estimates of target images  $I_1(x(\xi, T)^{-1})$  obtained with Tikhonov factor (a)  $\alpha = 0.5$  and (b)  $\alpha = 1$ , (c) Target image  $I_2(x(\xi), T)$ .

edges such as boundaries of a band of high concentration of Wingless protein observed in the ventral part of the wing disc.

All in all, the prior Gaussian model of the  $\lambda$ -field considered above suggests a typical growth pattern in the form of isolated dividing cells. Such a growth mode can be suitable for particular organisms but, certainly, it is not a universal growth mode. Since our main goal is not achieving a perfect image registration but rather revealing a qualitative picture of a biological growth via a natural measure of cell activities, namely, the Poisson intensity parameter  $|\lambda(\xi, t)|$ , we make an effort to find a more general prior model.

### VIII. The GRID intensity parameter $|\lambda(\xi, t)|$ as a random field governing the Poisson point process

We weaken the constraints on the  $\lambda$ -field and view it as a stochastic spatial process  $\{|\lambda(\xi_{seed})|, \xi_{seed} \in \Xi\}$  defined in a planar continuous domain  $\Xi$  at each time  $t$  of a realization of the macroscopic growth equation. Its absolute value influences the occurrences of cell divisions modeled as points of a spatial Poisson process (seeds). That is, points in  $\Xi$  are distributed according to a stationary Poisson process in time with the Poisson intensity  $|\lambda(\xi)|$ . The process points occur in  $\Xi$  with the following probabilities:

$$\Pr(N(\delta_{\xi_{seed}}) = 1 | \xi_s \in \Xi \setminus \delta_{\xi_{seed}}, \xi_s \neq \xi_{seed}) = \int_{\delta_{\xi_{seed}}} |\lambda(\xi)| d\xi + o(|\delta_{\xi_{seed}}|), \quad (\text{VIII.1})$$

where  $\delta_{\xi_{seed}} \in \Xi$  is a small region centered at the Poisson point  $\xi_{seed}$ ,  $N$  is a random number of seeds and  $\xi_s$  are all other seeds activated outside  $\delta_{\xi_{seed}}$ .

To be more specific, for a fixed time  $t$  we define a doubly stochastic counting Poisson point process  $\{N_\xi, \xi \in \Xi\}$  with intensity process  $\{|\lambda(\xi_{seed})|, \xi_{seed} \in \Xi\}$  in the Darcyan space  $\Xi$ ,

$$\Pr(N_A = n) = (n!)^{-1} \int_A |\lambda(\xi)| d\xi \cdot \exp\{-\int_A |\lambda(\xi)| d\xi\}, \quad (\text{VIII.2})$$

where  $A \subset \Xi$  and  $n$  is the number of activated seeds.

REMARK VIII.1. A spatial Poisson process is not totally separated from a temporal Poisson process. A stationary Poisson distribution in time is chosen as a model for

a spatial Poisson process observed at a particular time  $t = j$ . That is, the optimal  $|\lambda(\xi_{seed}, j)| = |\hat{\lambda}(\xi_{seed})|$ , where  $\xi_{seed}$  are activated seeds in  $\Xi$ -space with their total number  $N$ . At different times we have different observations of gene expression patterns leading to different MAP estimates of the underlying Poisson intensity of cell decisions. Therefore, a divorce of the spatial process from the temporal process is rather artificial.

Then at the time  $t$ ,  $\{x(\xi, t), \xi \in \Xi\}$  is a random two-dimensional vector field in the absolute space  $X$  influenced by the underlying doubly stochastic spatial Poisson point process. The Darcyan grid with  $n$  radial and  $m$  angular coordinate curves can be viewed as a result of the most likely realization of this process given observations in the form of images. It is the configuration of Darcyan nodes  $\{x(\xi_{1_i}, \xi_{2_j}, t), 1 \leq i \leq n, 1 \leq j \leq m\}$  constructed from the path of the process  $\{\xi : \xi \in \Xi\}$  yielding  $N = n \times m$  Poisson points  $\{\xi_{1_i}, \xi_{2_j}, t), 1 \leq i \leq n, 1 \leq j \leq m\}$  distributed with the optimal intensity parameter  $|\hat{\lambda}(\xi_{seed})|$ ,

$$x(\xi, t) = x(\xi, 0) + \sum_{j=1}^t \sum_{\xi_{seed}=1}^N \frac{\theta(x(\xi, (j-1)) - x(\xi_{seed}, (j-1)))}{J(x(\xi_{seed}, (j-1)))} \cdot \hat{\lambda}(\xi_{seed}), \quad (\text{VIII.3})$$

where  $\xi \in \{1, 2, \dots, N\}$ .

For the doubly stochastic Poisson process  $\{N_\xi, \xi \in \Xi\}$ , the probability density of finding  $N$  points in small neighbourhoods of  $\xi_1, \dots, \xi_N$  is

$$\begin{aligned} p[\xi_1 \in \delta(\xi_1), \xi_2 \in \delta(\xi_2), \dots, \xi_N \in \delta(\xi_N)] \\ = \prod_{[i=1]}^N \int_{\delta(\xi_i)} |\lambda(\xi)| d\xi \cdot \exp \left\{ - \int_{\delta(\xi_i)} |\lambda(\xi)| d\xi \right\}, \end{aligned} \quad (\text{VIII.4})$$

where the domain  $\Xi = \bigcup_{[i=1]}^N \delta(\xi_i)$ . Formula (VIII.4) defines the *sample-function density* for the spatial Poisson process with an arbitrary ordering of position.

Since the number of events is large, it is reasonable to assume that

$|\lambda(\xi)| = |\lambda(\xi_i)|$  in the small neighborhood of each Poisson point  $\xi_i$  of unit area,  $\delta_{\xi_i}$ , with  $1 \leq i \leq N$ . Then the sample-function density simplifies to

$$\begin{aligned} p_\xi = p[\xi_1 \in \delta(\xi_1), \xi_2 \in \delta(\xi_2), \dots, \xi_N \in \delta(\xi_N)] \\ = \prod_{[i=1]}^N |\lambda(\xi_i)| \cdot \exp\{-|\lambda(\xi_i)|\}. \end{aligned} \quad (\text{VIII.5})$$

The joint probability density for the doubly stochastic Poisson process appears to be a preferred representation of a prior knowledge about  $|\lambda(\xi, t)|$  emphasizing not only its space and time varying nature (at each time  $t$  we have a different realization of the  $|\lambda|$ -field) but also its role in governing a biological process of cell decisions. It is a general prior model since no specific probability measure has been assigned to a random  $|\lambda|$ -field. Having generated the initial Darcyan grid with  $N$  nodes, we view it as a result of a realization of a doubly stochastic Poisson spatial process. We find optimal measurements of the underlying Poisson intensity  $\{|\lambda(\xi)|\}$  at  $N$  seed locations  $\xi_i$  that provide the shortest path  $x(\xi, t)$  from the source image  $I_1(x(\xi))$  to the target image  $I_2(x(\xi), T)$ .

Using the negative of the sample-function density logarithm for the prior energy term in the cost function

$$E_{prior}(\lambda(\xi, t)) = -\ln p_{\xi, t} = -\sum_{[i=1]}^N \ln |\lambda(\xi_i, t)| + \sum_{[i=1]}^N |\lambda(\xi_i, t)|, \quad (\text{VIII.6})$$

we formulate the following optimal control problem:

$$\begin{aligned} \hat{\lambda}(\xi, T) = \arg \min_{\lambda \in R^N \times [0, T]} & \frac{1}{2\sigma_I^2} \sum_{\xi \in \Xi} (I_2(x(\xi), T) - I_1(x(\xi, T)^{-1}))^2 \cdot J(x(\xi)) \\ & - \sum_{[i=1]}^N \ln |\lambda(\xi_i, t)| + \sum_{[i=1]}^N |\lambda(\xi_i, t)|, \end{aligned} \quad (\text{VIII.7})$$

where  $x(\xi, t)$  is subject to the macroscopic growth equation (VI.6)

REMARK VIII.2. Since the cost function  $E_{post}$  depends on a high-dimensional intensity parameter  $\lambda(\xi, t)$ , it is not feasible to verify its convexity. It is possible that it possesses a few local minima and in this case a good initial guess on the  $\lambda$ -field would ensure convergence of the conjugate gradient algorithm to the closest local minimum. A natural choice for  $\lambda(\xi, 0)$  would be the uniform probability mass function assigned to  $N$  seeds in  $\Xi$ -space. This means that cell decisions are equally likely to occur anywhere in the organism's domain. We expect to obtain a perturbed version of the initially uniform distribution of seeds that would generate complex transformations reflecting inhomogeneous growth of the organism and bringing its source and target images into registration.

We now examine  $E_{prior}(\lambda(\xi, t)) : R^N \rightarrow R$  in more detail and drop the time index  $t$  for ease of notation. Its Hessian matrix,  $H(E_{prior})$ , is diagonal with entries

$$\frac{\partial^2 E_{prior}}{\partial^2 \lambda(\xi_i)} = \frac{1}{\lambda(\xi_i)^2}, \quad \frac{\partial^2 E_{prior}}{\partial \lambda(\xi_i) \partial \lambda(\xi_j)} = 0, \quad i \neq j, \quad (\text{VIII.8})$$

implying that it is positive-definite  $\forall \lambda \in R \setminus \{0\}^n$ ,

$$\lambda H(E_{prior}) \lambda^T \geq 0. \quad (\text{VIII.9})$$

The prior energy possesses a global minimum attained at the stationary vectors in  $R^N$  with components  $\lambda_s(\xi_i) = -1$  or  $\lambda_s(\xi_i) = 1$ . Indeed, if  $\lambda(\xi_i) > 0$ ,

$$\frac{\partial E_{prior}}{\partial \lambda(\xi_i)} = -\frac{1}{\lambda(\xi_i)} + 1 = 0 \Rightarrow \lambda_s(\xi_i) = 1.$$

Similarly, if  $\lambda(\xi_i) < 0$ ,

$$\frac{\partial E_{prior}}{\partial \lambda(\xi_i)} = -\frac{1}{\lambda(\xi_i)} - 1 = 0 \Rightarrow \lambda_s(\xi_i) = -1.$$

Thus we obtain the critical value of the Poisson intensity  $|\lambda_s(\xi)| = \bar{1}$  that maximizes the prior probability measure (VIII.5). The important implication of this result is that minimization of the prior energy term will force the Poisson intensity to stay close to the uniform probability distribution of seeds in  $\Xi$ -space at each time  $t$  and away from zero. This is consistent with the macroscopic growth law since it generates a growth deformation from a large number of seeds activated everywhere in the organism's domain. For an example of larval development of the *Drosophila* wing disc, such an optimal value

of the Poisson intensity reflects a biological evidence about the uniform division of cells throughout the whole cellular field.

The actual estimate of the growth parameter  $\hat{\lambda}(\xi)$  that is most likely to occur in the presence of observations  $(I_1(x(\xi)), I_2(x(\xi), T))$  provides more insight into the organism's growth. Due to its varying sign it identifies contraction and expansion subregions in the organism's interior caused by growth. Clearly, the likelihood term plays a dominant role in finding  $\hat{\lambda}(\xi)$  if there is a reason to believe that observed changes in pixel intensities are caused by a multitude of cell decisions such as cell division or death, for instance. In this case we would expect the estimated Poisson intensity  $|\hat{\lambda}(\xi)|$  to disperse from the uniform probability density of seed placements and thus show subregions of higher level of cell activities.

Minimization of the proposed posterior energy yields an estimate of the  $\lambda$ -field that does not preserve the diffeomorphic property of the transformation  $x(\xi, t)$ ,  $0 \leq t \leq T$  as suggested by computational experiments using a direct estimation algorithm. We increase smoothness properties of  $\lambda(\xi, t)$  by introducing an additional penalty term to our cost function in the form of the total variation integral,

$$\phi(\lambda) = \|\nabla \lambda\|_2^2 = \int_{\xi \in \Xi} |\nabla \lambda(\xi, t)|^2 d\xi,$$

as discussed in Section V. Including this term into the posterior energy and factoring out the coefficient  $\frac{1}{\sigma_I^2}$  we arrive at the following optimal control problem at a time  $t$ :

$$\begin{aligned} \hat{\lambda}(\xi, t) = \arg \min_{\lambda \in R^N \times [0, T]} & \sum_{\xi \in \Xi} [I_2(x(\xi), T) - I_1(x(\xi, t)^{-1})]^2 \cdot J(x(\xi)) \\ & + \sigma_I^2 \cdot \left[ \sum_{\xi_{seed} \in \Xi} [|\lambda(\xi, t)| - \ln |\lambda(\xi, t)|] + \sum_{\xi_{seed} \in \Xi} |\nabla \lambda(\xi, t)|^2 \right] \end{aligned} \quad (\text{VIII.10})$$

subject to the discretized macroscopic growth equation

$$\begin{aligned} x(\xi, t) = x(\xi, 0) + \sum_{j=1}^t \sum_{\xi_{seed} \in \Xi} [x(\xi, j-1) - x(\xi_{seed}, j-1)] \\ \cdot \exp\left(-\frac{\|x(\xi, j-1) - x(\xi_{seed}, j-1)\|^2}{step(x(\xi_{seed}, j-1))^2}\right) \cdot \frac{\lambda(\xi_{seed}, j)}{J(x(\xi_{seed}, j-1))} \end{aligned} \quad (\text{VIII.11})$$

for all seeds  $\xi = (\xi_1, \xi_2)$ ,  $1 \leq \xi_1 \leq n$ ,  $1 \leq \xi_2 \leq m$ .

Denote  $\sigma_I^2 = c_2$ . In the above setup,  $c_2$  plays a role of the weighting coefficient that determines the dominant term in the cost function. Before we apply the inference algorithm to our image data we have to specify the value of  $c_2$ . High values of  $c_2$  (of order  $O(1)$ ) make the prior energy a dominating term. This is undesirable since information given in observations becomes ignored. So, we find the value of  $c_2$  experimentally, starting from 1 and then decreasing its value by a factor of 0.1. The idea is to choose the value at which the posterior energy reaches the absolute minimum as a function of  $c_2$ .

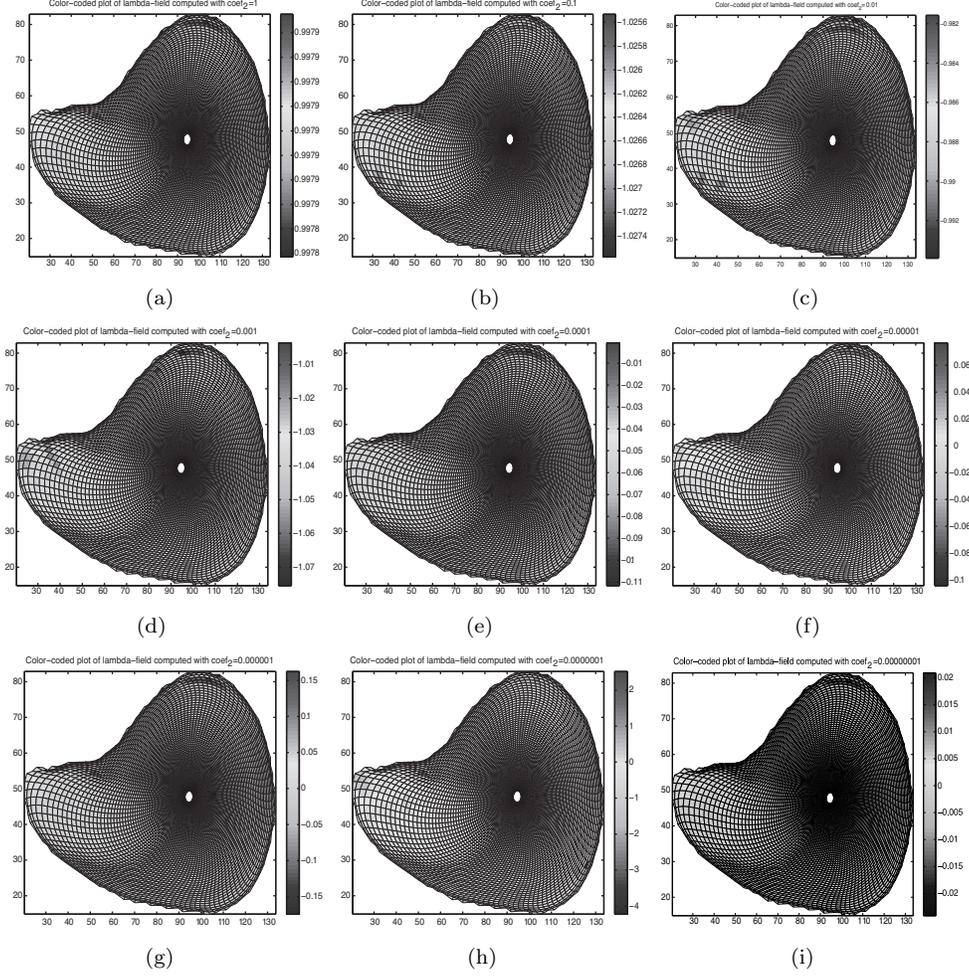


FIG. 9. (a-f) Color-coded plots of  $\hat{\lambda}(\xi)$  for the image pair  $(I_1, I_2)$  corresponding to the decreasing sequence of  $\{c_2(i)\}_{i=1}^9$ . In Fig. 9(a), the differences between the intermediate values are of order  $10^{-5}$ . The range is 0.9978–0.9979.

### IX. Performance of the image inference algorithm depending on the value of a prior energy coefficient

For a pair of images  $(I_1(x(\xi)), I_2(x(\xi), T))$  we perform 10 experiments with varying values of  $c_2$ . For each experiment  $i$ ,  $c_2(i) = 0.1^{i-1}$ . Shown in Figure 9 is the dynamics of  $\hat{\lambda}(\xi, T)$  as a function dependent on the value of  $c_2$ . For  $c_2 = 1$  the estimate  $\hat{\lambda}(\xi, T)$  lies nearly at the bottom floor of the prior energy surface valley given by  $\vec{1} \in R^N$  as seen from Figure 9(a). For the next value  $c_2 = 0.1$  the estimate of the  $\lambda$ -field falls into another valley of  $E_{prior}$ , namely,  $-\vec{1} \in R^N$  as seen from Figure 9(b). This is due to

the fact that both values of  $c_2$  make  $E_{prior}$  dominate over  $E_{likelihood}$ . Therefore, the obtained  $\lambda$ -estimates appear close to  $\vec{1}$  and  $-\vec{1}$ , the points of minimum of  $E_{prior}$  in  $R^N$ .

$\hat{\lambda}(\xi, T)$  remains negative for  $c_2$  of order between  $O(0.0001)$  and  $O(0.1)$  as shown in Figures 9(b)-9(e). This is not a meaningful estimate since it implies shrinking of the entire wing disc area. For  $c_2$  of order between  $O(0.1^8)$  and  $O(0.1^5)$ ,  $\hat{\lambda}(\xi, T)$  becomes mostly positive with higher values at the top edge of the anterior part of the wing disc shown in dark gray rapidly decreasing to negative values shown in black (see Figs. 9(f)-9(i)). This result suggests growth of the wing disc with a faster growing top edge causing contraction of the subregion lying next to it. Observe more complex behavior of  $\hat{\lambda}(\xi, T)$  with the presence of occasional highly localized expansions and a band of expansions running across the ventral part of the wing disc for a particular choice of  $c_2 = 0.1^7$  (see Figs. 9(h)-10(h)).

Shown in Figures 10(a)-10(i) are the deformed Darcyan grids generated by flows  $x(\xi, t)$ ,  $0 \leq t \leq T$  corresponding to the  $\lambda$ -estimates given in Figures 9(a)-9(i). As predicted by  $\hat{\lambda}(\xi, T)$ , the estimated transformation  $x(\xi, T)$  reflects pure growth for  $c_2 = 1$  (see Fig. 10(a)), pure decay for  $c_2$  of order between  $O(0.0001)$  and  $O(0.1)$  (see Figs. 10(b)-10(e)) with contractions mostly concentrated near the dorsal boundary of the organism, mixed development for  $c_2$  of order between  $O(0.1^8)$  and  $O(0.1^5)$  (see Figs. 10(f)-10(i)) with the combination of subregions of expansions and contractions mostly present near the top edge of the anterior part of the wing disc.

The magnitudes of the displacement  $\|x(\xi, 0) - x(\xi, T)\|$  appear concentrated near the dorsal boundary of the wing disc that diffuses into the interior for the  $c_2$  of order between  $O(0.0001)$  and  $O(1)$  as seen from Figures 11(a)-11(e). For lower orders of  $c_2$  the concentration becomes higher at the top edge of the anterior part (see Figs. 11(f)-11(i)) and more spread for  $c_2 = 0.1^7$  (see Fig. 11(h)).

All in all, the value of the prior energy weighting coefficient  $c_2$  influences qualitatively inference of the growth parameter  $\lambda(\xi, t)$ . As observed from Figure 9 meaningful estimates of the  $\lambda$ -field are obtained for orders of  $c_2$  lower than  $O(0.1^4)$ .

As the next step in our search for the ‘best’ estimate of the  $\lambda$ -field minimizing the posterior energy over the set  $\{1, 0.1, 0.1^2, \dots, 0.1^9\}$  of possible  $c_2$  values, we examine the posterior, prior and likelihood energies as functions of  $c_2$ . Recall that  $E_{post}(\lambda, c_2) = E_{likelihood}(\lambda) + c_2 \cdot [E_{prior}(\lambda) + \|\nabla \lambda\|_2^2]$ . Plots of these energies are presented in Figure 12 showing that the minimum is achieved at the value of  $c_2 = 0.1^7$ . Since the energy functions have been evaluated at a discrete set of points we can state that  $c_2 = 0.1^7$  is approximately a local minimizer of  $E_{prior}$  and an absolute minimizer of  $E_{likelihood}$  and  $E_{post}$ .

We conclude that with  $c_2 \sim O(0.1^7)$  we obtain a biologically meaningful estimate of the  $\lambda$ -field and the best estimate of the target image (see Fig. 13). As seen from Figure 9(h), the corresponding  $\lambda$ -field appears to be mostly positive throughout the domain of the wing disc except for a stripe-like subregion near the top edge of its anterior part shown in dark gray. Cell divisions have occurred everywhere with the peak of cell activities located near the top edge of the anterior part. This is a subregion with the highest mismatch of grey levels between source and target images  $I_1(x(\xi))$  and  $I_2(x(\xi), T)$  (see Figs. 13(a)-13(c)). Also, the corresponding growth deformation  $x(\xi, t)$  shown in

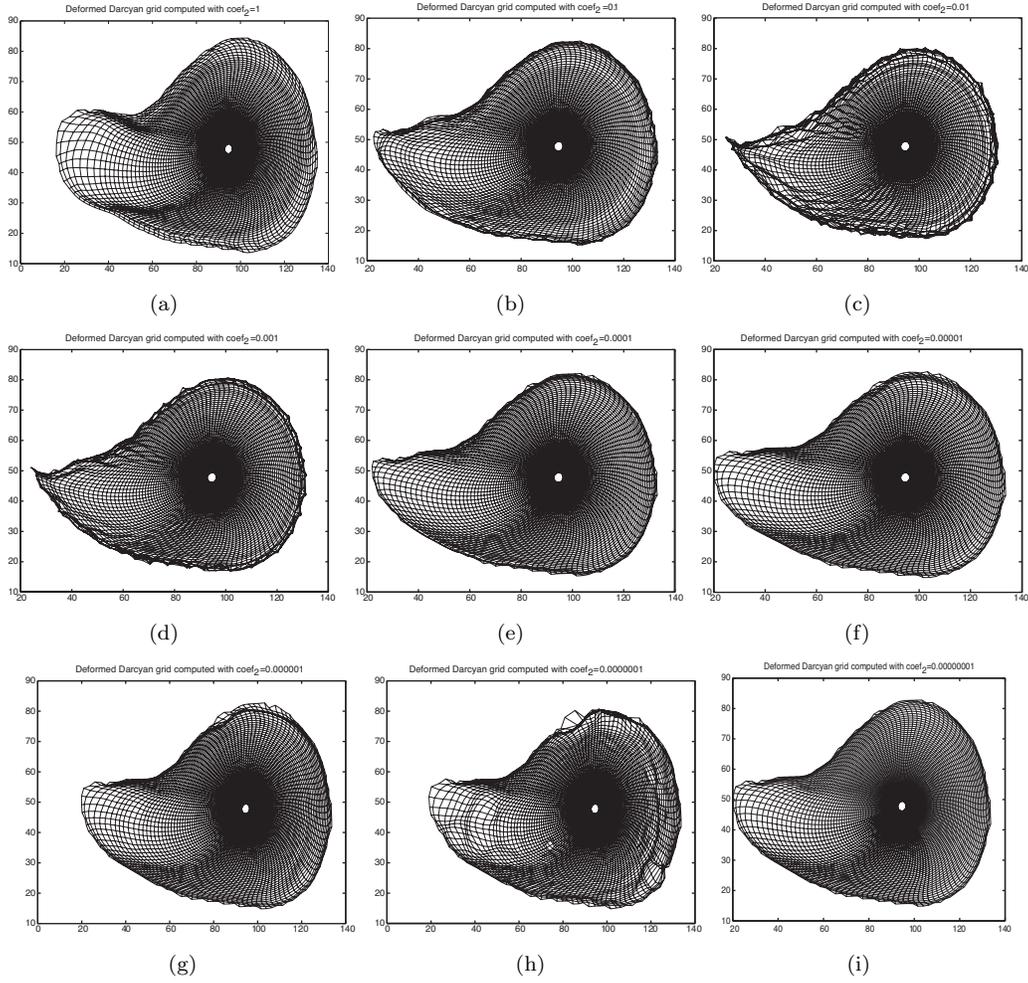


FIG. 10. (a-f) Deformed Darcyan grids of the wing disc generated by  $x(\xi, t)$ ,  $0 \leq t \leq T$  corresponding to the decreasing sequence of  $\{c_2(i)\}_{i=1}^9$ .

Figure 10(h) captures other changes in image pixel values such as a band of Wingless protein concentration near the ventral boundary seen in both images. It is possible that the rate of cell division is a bit higher in this subregion, thus causing local contractions of its neighborhood.

The prior model of the  $\lambda$ -field implemented above suggests a typical growth pattern in the form of cells dividing in clusters with the occasional appearance of isolated sites for the developing *Drosophila* wing disc. This result is qualitatively different from the one obtained using the prior Gaussian model. Elements of the estimated  $\hat{\lambda}(\xi, T)$  appear correlated with their neighbors due to a more general view of  $\lambda(\xi, t)$  as a random smooth field.

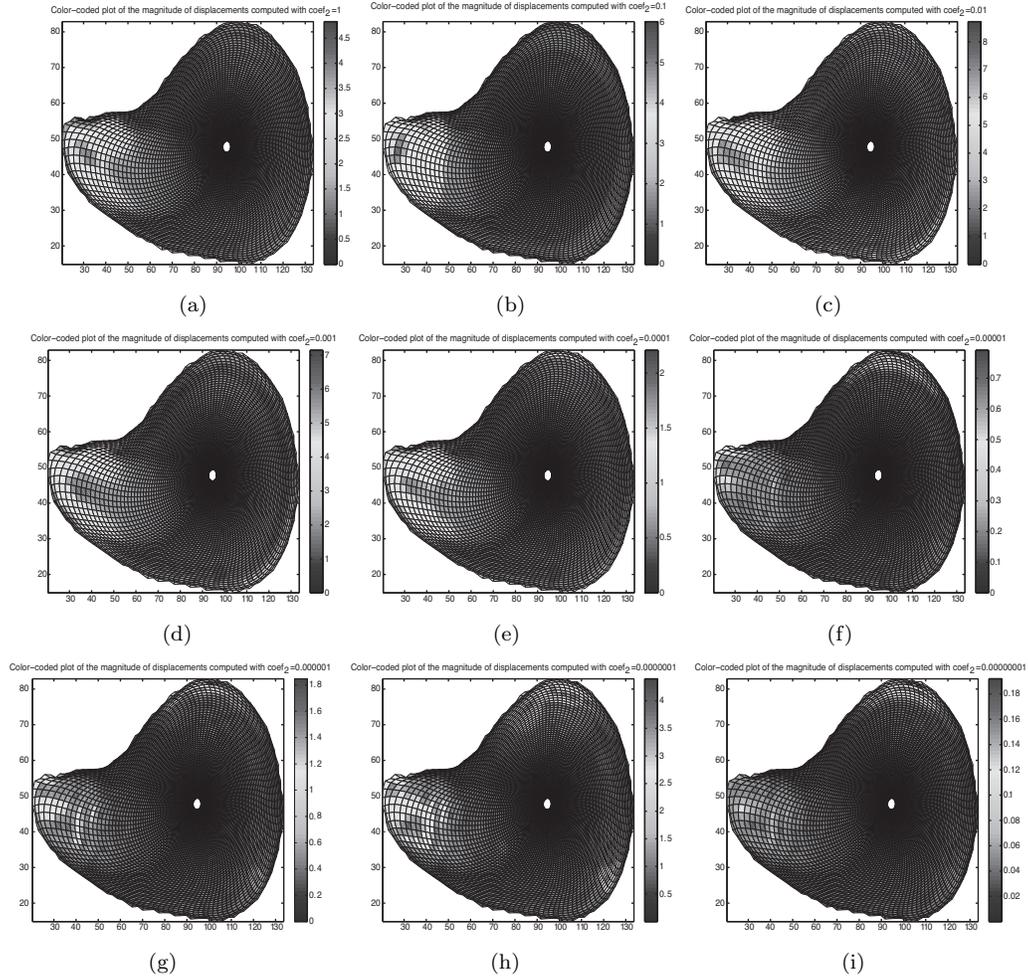


FIG. 11. (a-f) Magnitudes of the displacements corresponding to the decreasing sequence of  $\{c_2(i)\}_{i=1}^9$ .

Having experimental evaluation of the weighting coefficient of  $E_{prior}$  built into our inference algorithm we now infer growth properties of the *Drosophila* wing disc from the sequence of micrographs of a *Wingless* gene expression pattern at the larval stage of its development.

#### X. Revealing a hidden picture of growth from a sequence of micrographs

Taking into account the implicit dependency of the  $\lambda$ -field on the weighting coefficient of the prior energy term,  $c_2$ , we modify our original direct estimation algorithm to yield the ‘best’ estimate of  $\lambda(\xi, t)$  that corresponds to the value of  $c_2$  minimizing the cost function. We apply it to the consecutive pairs of images given in the sequence of micrographs of the *Wingless* expression pattern [18] in Figure 14. As a result, we reveal

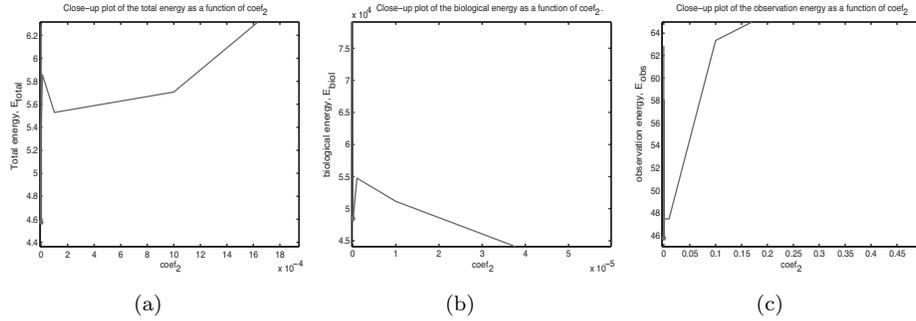


FIG. 12. Close-up plots of (a) the posterior energy  $E_{post}$ , (b) the prior energy  $E_{prior}$ , and (c) the likelihood energy  $E_{likelihood}$  as a function of  $c_2$ .

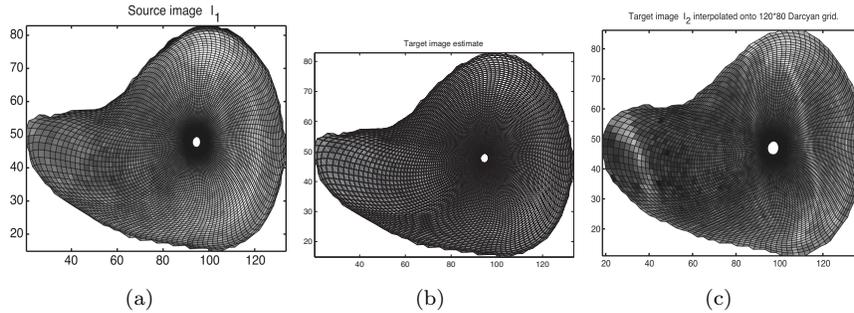


FIG. 13. (a) Source image  $I_1(x(\xi))$ , (b) Target image estimate  $I_1(x(\xi, T)^{-1})$ , (c) Target image  $I_2(x(\xi), T)$ .

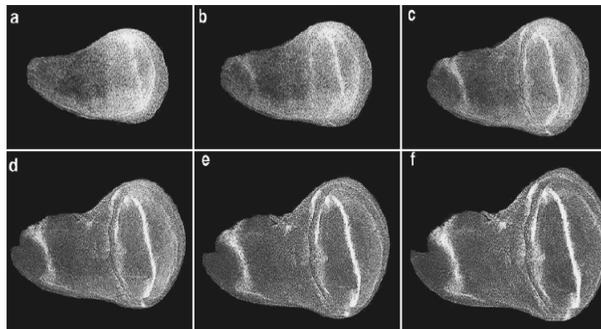


FIG. 14. Dynamics of Wingless gene expression pattern  $I(x, t)$  during the larval development of a wing disc from the second ( $\approx 48$  hours)(a) through the third instar ( $\approx 120$  hours)(f)

{Courtesy of S. Paddock and D. McDougal, University of Wisconsin, USA}

the underlying dynamics of the  $\lambda$ -field controlling the growth-induced deformation of the Darcy grid of the *Drosophila* wing disc  $x(\xi, t)$  as shown in Figures 15(b), 15(e), 15(h), 15(k), 15(n). At the early second instar of larval development (see Figs. 15(a)-15(c)) the

subregion of high cell activities stretches near the top edge of the anterior part where the organism undergoes large contractions next to the thin band of expansions. The magnitudes of displacement form patterns of oval rings with the largest values located at the top anterior edge.

There is also a visible concentration of lesser-magnitude deformations in the dorsal part of the wing disc. Mostly positive  $\lambda$ -values here suggest a faster growth of this region as compared to the central part of the disc.

As seen from Figures 15(d)-15(f) the peak of high cell activities moves to the dorsal part of the wing disc. Faster growth of the dorsal part occurs in a pattern of a large and two small discs with the largest magnitude of displacements forming a wide concentric ring. There is still a presence of large contractions close to the top edge of the anterior part, but the magnitudes-of-displacements pattern is changed to a couple of separate band-like subregions.

Figures 15(g)-15(i) show a high concentration of cell divisions near the tip of the dorsal part causing contractions of larger/smaller magnitude in the anterior/posterior directions. The organism also undergoes contractions near the top anterior edge, but the magnitude of this deformation is significantly decreased.

In the following triplet of Figures 15(j)-15(l), subregions of contraction appear larger in size and with smaller magnitudes of deformation in the anterior direction of dorsal and posterior parts and near the top anterior boundary. The intensity of cell divisions is relatively uniform throughout the wing disc. As a result, the magnitudes-of-deformations pattern evolves in the form of a larger disc in the dorsal part and band-like subregions in the posterior part and near the top anterior boundary.

At the late third instar of larval development seen in Figures 15(m)-15(o), subregions of contraction become more spread throughout the wing disc with the largest one arising as a wide band across the dorsal part. There is also a band-like subregion of expansions with contractions on both sides seen in the ventral part of the disc. Such deformations capture changes in the location of the highest concentration of the Wingless protein. A combination of local contractions and expansions yields a complex growth deformation manifesting itself in a variety of forms throughout the wing disc.

The modified direct estimation algorithm based on the prior energy in the form of a sample-function density allows us to capture local deformations inside the domain of the organism. As seen from Figures 16(a) and 16(c), a band of the highest concentration of the Wingless protein becomes slightly wider. One can imagine a series of cell divisions occurring in this narrow subdivision of the wing disc. The deformation effects of such cell activities are shown as the expansion band with contractions on both sides in the deformed Darcyan grid.

## XI. Summary and conclusions

In this paper, we have addressed the inference problem of growth properties of an organism directly from image data expressed in terms of GRID variables, namely, the Poisson intensity  $|\lambda(\xi, t)|$  and the macroscopic biological transformation  $x(\xi, t)$ . *Maximum a posteriori* estimation of the GRID growth parameters has been performed for a

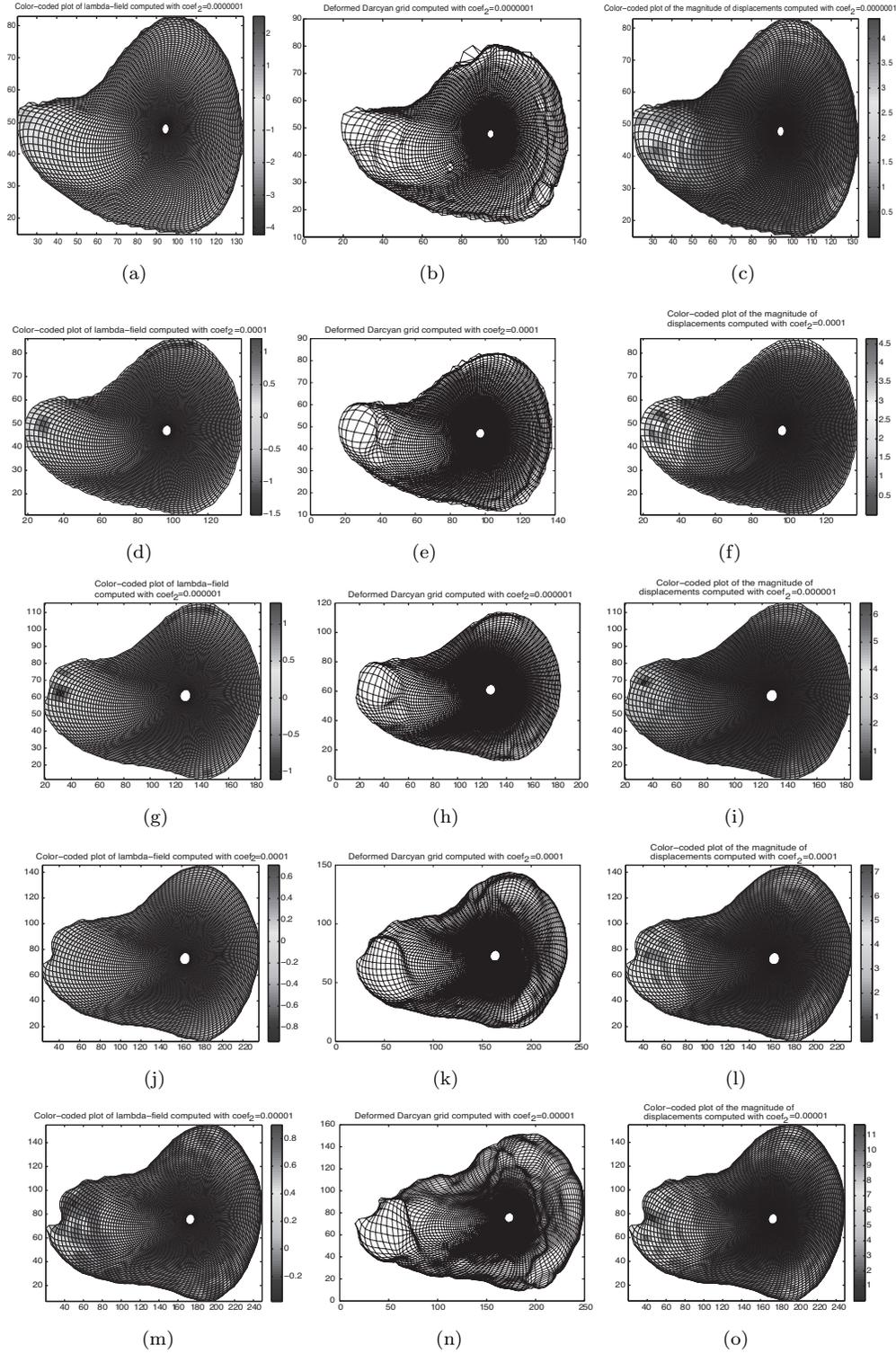


FIG. 15. The best estimates of  $\lambda(\xi, T)$ ,  $x(\xi, T)$  and  $\|\Delta x(\xi, T)\|$  for an image pair (a-c)  $I_1(x(\xi)), I_2(x(\xi), T)$ ,  $c_2 = 0.1^7$ , (d-f)  $I_2(x(\xi)), I_3(x(\xi), T)$ ,  $c_2 = 0.1^4$ , (g-i)  $I_3(x(\xi)), I_4(x(\xi), T)$ ,  $c_2 = 0.1^6$ , (j-l)  $I_4(x(\xi)), I_5(x(\xi), T)$ ,  $c_2 = 0.1^4$ , (m-o)  $I_5(x(\xi)), I_6(x(\xi), T)$ ,  $c_2 = 0.1^5$ .

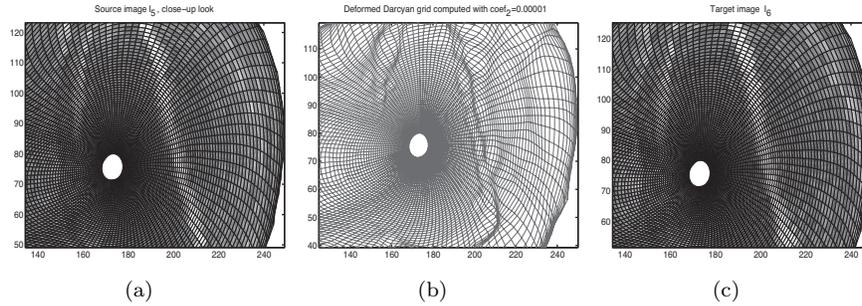


FIG. 16. Close-up fragment of (a) the source image  $I_5(x(\xi))$ , (b) the deformed Darcyan grid  $(x(\xi, T))$  and (c) the target image  $I_6(x(\xi, T))$ .

sequence of micrographs of the larval *Drosophila* wing disc using the GRID macroscopic growth law.

Firstly, we have studied the GRID macroscopic growth integro-differential equation. It models the transformation of the source image of a specimen into its target image by a diffeomorphic flow dependent on  $\lambda(\xi, t)$  whose absolute value represents the Poisson intensity of cell decisions. The discrete version of the GRID macroscopic growth law guarantees invariance of the inference  $\hat{\lambda}(\xi, t)$  under the Euclidean group of transformations and its numerical instability due to the ill-conditioned matrix representation of the macroscopic growth operator. As such, remedies in the form of Tikhonov regularization and an additional constraint on  $\lambda(\xi, t)$  to be a smooth function in the organism's domain have been proposed and implemented.

Secondly, using a Bayesian interpretation of the unknown  $\lambda(\xi, t)$ , we have explored such prior models of  $\lambda$ -parameter as the Gaussian random field with i.i.d. elements and a random smooth field governing the inhomogeneous Poisson process of cell decisions. The latter prior model defined by the sample-function density for a doubly stochastic Poisson point process appears to be more natural since no particular probability measure has been assigned to the random  $\lambda$ -field. The important advantage of this model over the former one is that the *maximum a posteriori* estimate of  $\lambda(\xi, t)$  contains stochastically dependent elements with neighborhoods of influence of varying size. In the Gaussian prior model, the size of the neighborhoods of influence is restricted by the prescribed constant variance  $\sigma_\lambda$  resulting in the oscillating behavior of  $\lambda(\xi, t)$ .

Therefore, the cost function modeled as the posterior energy in the form of the sum of the likelihood and the preferred prior energies leads to more complex, and thus more realistic, optimal growth patterns. They drive local changes in pixel intensities towards a registration of images provided that the likelihood energy dominates over the prior energy. The dominance is controlled by the weighting coefficient of the prior energy term  $c_2$ . We have found experimentally that for a certain range of orders of magnitude of  $c_2$ , the inference algorithm yields biologically meaningful estimates of the  $\lambda$ -field. The 'best'  $\lambda$ -estimate corresponds to the value of  $c_2$  at which the posterior energy reaches the absolute minimum as a discrete function of  $c_2$  defined at  $\{0.1^{i-1}\}_{i=1}^{10}$ .

Since optimization of the preferred cost function does not preserve the topology of a growing organism, we have included a penalty term in the form of the total variation integral. It generates smooth  $\lambda$ -estimates; however, in areas of high grey-level mismatch, drastic changes in  $\lambda$ -values may cause an overlap of the deformed Darcyan grid. Gradient and Laplacian penalty functions (the first and second derivatives of  $\lambda(\xi, t)$ ) can be used to increase the degree of smoothing.

The diffeomorphic flow corresponding to the optimal value of the  $\lambda$ -field does not fully register images of the Wingless gene expression patterns of the growing wing disc. This is due to the smoothness constraints imposed on  $\lambda(\xi)$  to ensure that the integrity of the Darcyan grid under the growth deformation is preserved.

More realistic MAP estimates of GRID growth variables can be obtained for the larval *Drosophila* wing disc by including more biological data into the cost function model. We have only considered micrographs of Wingless gene expression patterns. There are other key regulatory genes, namely, Vestigial and Apterous, required for the wing formation in *Drosophila* during its larval development whose spatial-temporal patterns of expression may depend on cell division process. If we had micrographs of their expression patterns we would modify the posterior energy as follows:

$$E_{post} = E_{obsWingless} + E_{obsVestigial} + E_{obsApterous} + c_2 \cdot [E_{prior} + \|\nabla\lambda\|_2^2].$$

All in all, the proposed cost function model that measures cell activities represented by the Poisson intensity of cell decisions  $|\lambda(\xi, t)|$  has potential applications in the analysis of micrographs of gene expression patterns in developing organisms. The micrographs provide direct biological data since pixel intensities represent concentrations of a certain gene product distributed within the organism's domain. If there is a biological reason to believe that the dynamics of gene expression patterns is correlated with an underlying biological process of cell decisions, then hidden growth patterns reflecting gross anatomical changes as controlled by the Poisson intensity parameter can be revealed.

#### Appendix A. Direct estimation algorithm of the GRID growth parameter $\lambda(\xi, t)$ and the growth-induced transformation $x(\xi, t)$

In order to solve the discrete optimal control problem for  $\hat{\lambda}(\xi, T)$  we implement the following steps of the estimation algorithm:

- (1) Construction of the Darcyan coordinate system of the wing disc seen in source  $I_1$  and target  $I_2$  images using the Level Set Method as described in [14].

Since there is no difference in the grey level contents of both images in the neighborhood of the Darcyan origin, we stop propagating the level sets into this area. The resulting Darcyan grid contains 80 radial and 120 angular coordinate lines. Since the image function  $I$  is defined on a rectangular uniform lattice we apply bilinear interpolation to estimate image values at the nodes of the curvilinear Darcyan coordinate system  $I(x(\xi_i))$ , where  $1 \leq i \leq N$  and  $N = 80 \times 120$ .

- (2) Image preprocessing.

- (i) Image registration about the origin of the Darcyan coordinate system applying Principal Component Analysis (PCA). We remove the rigid transformations from observed images using Darcyan coordinate representations of the initial and grown wing discs. Given the Darcyan grid nodes  $x_i = x(\xi_{1_i}, \xi_{2_i})$ ,  $1 \leq i \leq N$ , we define the covariance matrix  $K = \frac{1}{N} \sum_{n=1}^N \{(x_i - x_c)(x_i - x_c)^T\}$  and its eigenvectors for  $I_1$  and  $I_2$ .  $x_c$  is the origin of the Darcyan coordinate system of the grown wing disc shown in  $I_2$ .
- The eigenvectors of  $K$  give the principal axes of both Darcyan. Then we register  $I_1$  with  $I_2$  by aligning the principal axis of their Darcyan. That is,  $x_{i_{I_1}(reg)} = R x_{i_{I_1}} + t$ , where  $R$  is the rotation matrix and  $t$  is the translation vector.
- (ii) Salt and pepper noise removal. Salt and pepper noise is inherent to the confocal microscopy imaging technique. A median filter using a  $3 \times 3$  window can be applied as it has a property of removing this kind of noise while preserving edges.
- (3) Realization of the conjugate gradient method to find the optimal value of the growth magnitude  $\hat{\lambda}(\xi)$ .

We apply the Polak-Ribière version [19] since it is useful in finding the minimum of a nonconvex function in  $R^N$ : the repeated cycles of  $N$  iterations eventually converge to a minimum. We construct a minimizing sequence of optimal controls  $\{\lambda(\xi, t)\}_{t=1}^T$  such that  $\lim_{t \rightarrow T} \lambda(\xi, t) = \hat{\lambda}(\xi)$  and  $I_2(x(\xi), T) \approx I_1(x(\xi, T)^{-1})$ .

At each iteration  $t$  we compute the energy gradients

$$\begin{aligned} \nabla E_{likelihood}(\lambda(\xi, t)) &= -2 \sum_{\xi_j \in \Xi} (I_2(x(\xi_j), T) - I_1(x(\xi_j, t)^{-1})) \\ &\quad \cdot \left[ \frac{\partial I_1(x(\xi_j, t)^{-1})}{\partial x_1(\xi_j)} \cdot \frac{\partial x_1(\xi_j)}{\partial \lambda} + \frac{\partial I_1(x(\xi_j, t)^{-1})}{\partial x_2(\xi_j)} \cdot \frac{\partial x_2(\xi_j)}{\partial \lambda} \right] \cdot \left| \frac{\partial(x_1, x_2)}{\partial(\xi_{1_j}, \xi_{2_j})} \right|, \\ \nabla E_{prior}(\lambda(\xi, t)) &= -\frac{1}{\lambda(\xi, t)} + 1, \quad \lambda(\xi, t) > 0, \\ \nabla E_{prior}(\lambda(\xi, t)) &= \frac{1}{\lambda(\xi, t)} - 1, \quad \lambda(\xi, t) < 0, \end{aligned}$$

where  $x(\xi_j) = (x_1(\xi_j), x_2(\xi_j))$ ,  $1 \leq \xi_j \leq N$  and the gradient of the penalty  $\sum_{[i=1]}^N |\lambda(\xi_i, t)|$ . The Darcyan grid is updated as

$$x(\xi, t) = x(\xi, t-1) + \sum_{\xi_{seed_j} \in \Xi} \theta(x(\xi_i, t-1) - x(\xi_{seed_j}, t-1)) \cdot \lambda(\xi_{seed_j}, t). \quad (\text{A.1})$$

*Computational aspects of the Polak-Ribière algorithm.* At each iteration of the Polak-Ribière procedure we need to minimize the energy functional  $E_{post}(\lambda(\xi, t))$  along the direction  $h_i \in R^N$  conjugate to all previous directions. For this purpose we solve the one-dimensional line minimization problem

$$a_t = \arg \min_{a \geq 0} \{E_{post}(\lambda(\xi, t) + a \cdot h_i(\xi))\}$$

that appears to be nontrivial when the behavior of  $E_{post}(\lambda(\xi, t_i))$  is unknown. It is tempting to approximate the energy functional as a function of  $\lambda$  by a high-order interpolating polynomial. Indeed, it is possible to construct it based on the information of the energy functional and its derivative values at the endpoints of the interval  $[0, b]$  where the unknown  $a$  belongs. This allows us to avoid evaluations of the image function at each updated value of  $a_t$  and save computational time. However, this approach does not guarantee right interpolation especially if  $E_{post}(\lambda(\xi, t))$  has sharp features. Also, it is not clear how far we should go in search for the minimum, that is, what value of  $b$  we should choose.

We use a Golden Section-based algorithm of one-dimensional minimization [20] that first brackets the local minimum and then isolates it using the signs of the energy derivative on both sides of the minimum. The return of the zeroth value of  $a$  by this code is used as a stopping criterion for the multidimensional minimization. Indeed, in this case,  $E_{post}$  attains its minimum at the previous value of  $a_t$ .

Implementing such a 1D minimization routine, we certainly lose in computational time but we gain in precision of the image estimates!

#### REFERENCES

- [1] U. Grenander, M. I. Miller: *Computational anatomy: An emerging discipline*. Quart. Appl. Math. **56(4)** 617–694 (1998) MR1668732 (2000e:92011)
- [2] F. L. Bookstein: *The measurement of biological shape and shape change*. Lecture Notes in Biomathematics, Springer-Verlag (1978)
- [3] F. L. Bookstein: *Morphometric tools for landmark data: Geometry and biology*. Cambridge University Press (1991) MR1469220 (99d:92003)
- [4] F. L. Bookstein: *Biometrics, biomathematics and the morphometric synthesis*. Bulletin of Mathematical Biology **58(2)** 313–365 (1996)
- [5] C. Davatzikos: *Spatial transformation and registration of brain images using elastically deformable models*. Comput. Vision Image Understanding **66(2)** 207–222 (1997)
- [6] S. K. Kyriacou, C. Davatzikos, S. J. Zinreich, R. N. Bryan: *Nonlinear elastic registration of brain images with tumor pathology using biomechanical model*. IEEE Trans. Med. Imaging **18(7)** 580–592 (1999)
- [7] G. E. Christensen, R. D. Rabbitt, M. I. Miller: *A deformable neuroanatomy textbook based on viscous fluid mechanics*. Proceedings of 1993 Conference on Information Sciences and Systems, Johns Hopkins University 211–216 (1993)
- [8] G. E. Christensen, R. D. Rabbitt, and M. I. Miller: *Deformable templates using large deformation kinematics*. IEEE Transactions on Image Processing, **5(10)** 1433–1447 (1996)
- [9] G. E. Christensen, C. J. Sarang, M. I. Miller: *Volumetric Transformation of Brain Anatomy*. IEEE Transactions on Med. Imaging **16(6)** 864–876 (1997)
- [10] A. Srivastava, S. Saini, Z. Ding, U. Grenander: *Maximum-likelihood Estimation of Biological Growth Variables*. Energy Minimization Methods in Computer Vision and Pattern Recognition, LNCS**3757** 107–118 (2005)
- [11] U. Grenander: *On the Mathematics of Growth*. Quart. Appl. Math. **65** 205–257 (2007) MR2330557 (2008e:92007)
- [12] U. Grenander, A. Srivastava, S. Saini: *A Pattern-theoretic Characterization of Biological Growth*. IEEE Trans. Med. Imaging **26(5)** 648–659 (2007)
- [13] N. Portman, U. Grenander, E. R. Vrscay: *Direct Estimation of Biological Growth Properties from Image Data Using the “GRID” Model*. To be published in Proceedings of the 6th International Conference on Image Analysis and Recognition (ICIAR) (2009)
- [14] N. Portman, U. Grenander, E. R. Vrscay: *New Computational Methods of Construction of Darcyan Biological Coordinate Systems*. Image Analysis and Recognition, LNCS**4633** 143–156 (2007)

- [15] S. B. Carroll: *Endless Forms Most Beautiful. The New Science of Evo Devo and the Making of Animal Kingdom*. W. W. Norton and Company Inc., NY (2005)
- [16] J. A. Williams, S. W. Paddock and S. B. Carroll: *Pattern Formation in a Secondary Field: a Hierarchy of Regulatory Genes Subdivides the Developing Drosophila Wing Disc into Discrete Sub-regions*. *Development* **117** 571–584 (1993)
- [17] U. Grenander and M. Miller: *Pattern Theory: From Representation to Inference*. Oxford University Press, Inc., NY (2007) MR2285439 (2008d:94003)
- [18] S. Paddock and D. McDougal: *Making Movies on a Weekend*. *BioImaging* **29(5)** 997–1004 (2000)
- [19] E. Polak: *Computational Methods in Optimization: a Unified Approach*. Mathematics in Science and Engineering, Academic Press, NY **77** (1971) MR0282511 (43:8222)
- [20] W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling: *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press (1992) MR1201159 (93i:65001b)