Proceedings of the ASME 2018 Dynamic Systems and Control Conference
DSCC2018
September 30-October 3, 2018, Atlanta, Georgia, USA

DSCC2018-9164

MODELING OF COLLECTIVE CELL BEHAVIORS INTERACTING WITH EXTRACELLULAR MATRIX USING DUAL FACETED LINEARIZATION

Michaëlle N Mayalu
Division of Engineering and Applied Science,
California Institute of Technology
Pasadena, CA ,USA

Min-Cheol Kim
Department of Mechanical Engineering,
Massachusetts Institute of Technology
Cambridge, MA ,USA

Haruhiko Asada
Department of Mechanical
Engineering, Massachusetts
Institute of Technology
Cambridge, MA ,USA

ABSTRACT

Cells interacting over an extracellular matrix (ECM) exhibit emergent behaviors, which are often observably different from single-cell dynamics. Fibroblasts embedded in a 3-D ECM, for example, compact the surrounding gel and generate an anisotropic strain field, which cannot be observed in single cell-induced gel compaction. This emergent matrix behavior results from collective intracellular mechanical interaction and is crucial to explain the large deformations and mechanical tensions that occur during embryogenesis, tissue development and wound healing. Prediction of multi-cellular interactions entails nonlinear dynamic simulation, which is prohibitively complex to compute using first principles especially as the number of cells increase. Here, we introduce a new methodology for predicting nonlinear behaviors of multiple cells interacting mechanically through a 3D ECM. In the proposed method, we first apply Dual-Faceted Linearization to nonlinear dynamic systems describing cell/matrix behavior. Using this unique linearization method, the original nonlinear state equations can be expressed with a pair of linear dynamic equations by augmenting the independent state variables with auxiliary variables which are nonlinearly dependent on the original states. Furthermore, we can find a reduced order latent space representation of the dynamic equations by orthogonal projection onto the basis of a lower dimensional linear manifold within the augmented variable space. Once converted to latent variable equations, we superpose multiple dynamic systems to predict their collective behaviors. The method is computationally efficient and accurate as demonstrated through its application for prediction of emergent cell induced ECM compaction.

INTRODUCTION

The compaction of a cell-populated fibrous extracellular matrix (ECM) is an important mechanism for numerous biological processes including embryogenesis [1], tissue development [2] and wound healing [3–5]. In addition, during wound healing traction forces exerted by fibroblasts and myofibroblasts result in ECM compaction at the site of injury [3–5].

ECM gel compaction is a complex process from the mechanics point of view. The fibrous nature of the extracellular matrix (ECM) that forms a network of cross-linked fibers is highly nonlinear and intricate, but is critical for predicting large compaction and long-range transmission of forces [6]. As a large deformation is induced, the standard linear mechanics model yields significant errors since the ECM fiber network is anisotropic and even causes irreversible deformations as a large compaction takes place. This prominent nonlinearity prohibits us from using simple methods for predicting the ECM compaction by a multitude of cells. Contributions of individual cells cannot simply be added, or superposed, to predict a resultant compaction of the fibrous ECM. Cells' properties, too, are highly nonlinear and complex. Considering these nonlinear physical and physiological properties results in a complex computational model consisting of differential equations that are intractably complex due to high-dimensional, nonlinear coupled dynamics.

Previous in-silico models predicting compaction of cell-populated gels have been used to supplement and enhance in-vitro experiments and allow for fast and inexpensive methods to evaluate various cell types and environmental conditions [7],

[8–13]. In addition computational models can help elucidate and explore the underlying phenomena involved in compaction.

Although mathematical continuum [7], [12], [11]and finite-element models[8] of cell-mediated gel compaction exist, agent-based modeling has been more recently used to study emergent phenomena because it can predict the adaptive behavior of individual components as a result of underlying rules [9,10]. However, current agent-based models often require abstraction of details in order to limit mathematical and computational complexity and computational cost especially for modeling larger cell populations. Most significant abstractions identified within [7], [8–13] are (1) 2-D representation of a 3-D system; (2) exclusion of intracellular mechanics (3) absence of realistic ECM fiber mechanics. In addition, simple rules describing the interactions between agents are often determined somewhat heuristically. These simplifications may put limitations to the model validity and furthermore the translational potential.

Using Dual Faceted Linearization, the original nonlinear dynamics of cells and ECM derived from physical and physiological principles are recast in an enlarged state space by augmenting independent state variables with auxiliary variables that inform all the nonlinear forces and displacements involved in the system. Once represented in the augmented space, the cells and ECM can be treated as linear systems which facilitates modular computation, separation and division of the coupled system, and scalable and integrative analysis. The linear representation further facilitates latent variable transformation and model order reduction. Within the linear latent variable representation, collective cell behaviors can be predicted by merely superposing the contributions of the individual cells. Using this method, computational expense and time are decreased significantly and sufficient mechanistic detail is retained in the simulation.

NONLINEAR GOVERNING EQUATIONS FOR COLLECTIVE CELL BEHAVIORS IN ECM FIBER NETWORK

Consider multiple cells having an identical phenotype are interacting to each other through the surrounding 3D ECM fiber network. We construct a computational model for predicting cell-mediated ECM gel compaction. ECM is modeled as a network of many fibers connected at a large number of nodes (N_e ≈2000), whereas each cell is represented with a mesh structure consisting of many nodes ($N_c \approx 200$). Each cell binds to the surrounding ECM fiber network by forming focal adhesions. Acting on the i-th node of the k-th cell with spatial coordinates $\boldsymbol{x}_i^{c,k} \in \Re^{3\times 1}$ are the cell's cortical tension force and elastic energy force (collectively denoted as $\mathbf{F}_{Cort-Elas,i}^{c,k} \in \Re^{3\times 1}$) focal adhesion force (denoted as $\mathbf{F}_{FA,i}^{c,k} \in \Re^{3\times 1}$), lamellipodium force ($\mathbf{F}_{L,i}^{c,k} \in \Re^{3\times 1}$) and frictional damping force ($\mathbf{F}_{Damp,i}^{c,k} \in \Re^{3\times 1}$). Assuming that the mass of the node is negligibly small and the damping force is given by $\mathbf{F}_{Damp,i}^{c,k} = -D\, d\boldsymbol{x}_i^{c,k}/dt$, where D is damping constant, the equation of motion is given by:

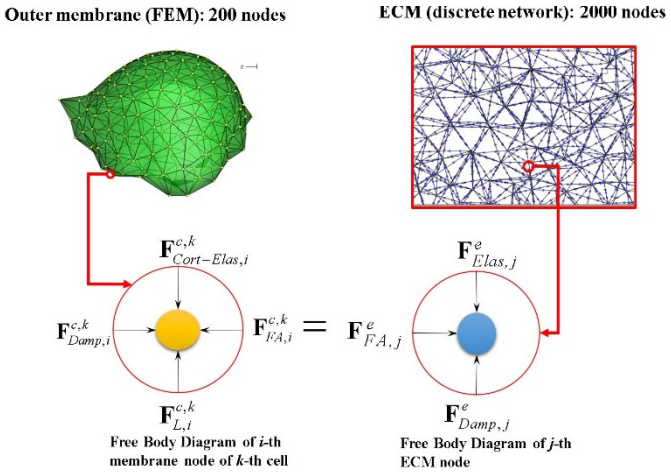

FIGURE 1: SCHEMATIC DIAGRAM OF CELL-ECM INTERACTION

$$\mathbf{F}_{Cort-Elas,i}^{c,k} + \mathbf{F}_{FA,i}^{c,k} + \mathbf{F}_{L,i}^{c,k} - D\frac{d\boldsymbol{x}_i^{c,k}}{dt} = 0, \quad i = 1,\cdots,N_c, k = 1,\cdots,n_{cell} \quad (1)$$

The cortical tension and elastic energy force of the k-th cell is a nonlinear function of its membrane coordinates. The focal adhesion force can be approximated to a nonlinear algebraic function of cell membrane and ECM nodes as well as the biochemical parameters involved in integrin-ligand binding. These nonlinear relationships may be found in [14,15]

The forces acting on the j-th node of the fiber network include the elastic energy forces (including both lateral restoring forces and bending moments, $\mathbf{F}_{Elas,j}^{e}$), focal adhesion forces (from the shared attachment with the cell, $\mathbf{F}_{FA,j}^{e}$) and damping forces ($\mathbf{F}_{Damp,j}^{e}$) [*25, 26*]. The equation of motion can be written as:

$$\mathbf{F}_{Elas,j}^{e} + \mathbf{F}_{FA,j}^{e} - D_e\frac{d\boldsymbol{x}_j^{e}}{dt} = 0 \quad , j = 1,\cdots,N_e \quad (2)$$

The ECM elastic energy force is a nonlinear function of ECM coordinates [14,15]. Physics dictates that the focal adhesion force of the i-th membrane node of the k-th cell attached to the j-th ECM node is:

$$\mathbf{F}_{FA,i}^{c,k} = -\mathbf{F}_{FA,j}^{e} \quad (3)$$

Assuming that no two cells bind to the same ECM node, $\mathbf{F}_{FA,i}^{c,k}$ and $\mathbf{F}_{FA,j}^{e}$ have the same magnitude with the opposite signs. The focal adhesion connections between the membrane nodes and ECM nodes change over time as the cell membrane deforms, gains traction and generates lamellipodial protrusions.

Although the governing equations derived above are rigorous and based on basic principles, they are complex and can become computationally expensive as the number of cells increase. The number of state variables for the given system is $3N_e+3N_cn_{cell}$, which is on the order of 7,000 for $n_{cell}=2$. We aim to a) linearize the system using Dual faceted Linearization b) considerably reduce the number of state variables though latent variable transformation and c) predict collective behaviors of the multiple cells through superposition of individual cell dynamics.

DUAL-FACETED LINEARIZATION

In Dual Faceted Linearization, we represent the nonlinear dynamical system in an augmented space consisting of independent state variables and nonlinear forces as the additional variables, termed auxiliary variables. Let us define the augmented space with two sets of linear differential equations:

$$\left.\begin{aligned}\frac{d\boldsymbol{x}^{c,k}}{dt} &= \mathbf{W}^{c}_{CE}\mathbf{F}^{c,k}_{Cort-Elas} + \mathbf{W}^{c}_{FA}\mathbf{F}^{c,k}_{FA} + \mathbf{L}_c\mathbf{u}^k, \quad k=1,\cdots,n_{cell}\\ \frac{d\boldsymbol{x}^{e}}{dt} &= \mathbf{W}^{e}_{Elas}\mathbf{F}^{e}_{Elas} + \mathbf{W}^{e}_{FA}\mathbf{F}^{e}_{FA}\end{aligned}\right\} set\ 1 \quad (4)$$

$$\left.\begin{aligned}\frac{d\mathbf{F}^{c,k}_{Cort-Elas}}{dt} &\simeq \mathbf{Q}^{c}_{x}\boldsymbol{x}^{c,k} + \mathbf{Q}^{c}_{F_{CE}}\mathbf{F}^{c,k}_{Cort-Elas} + \mathbf{Q}_u\mathbf{u}^k, \quad k=1,\cdots,n_{cell}\\ \frac{d\mathbf{F}^{c,k}_{FA}}{dt} &\simeq \mathbf{H}^{c}_{x}\boldsymbol{x}^{c,k} + \mathbf{H}^{e}_{x}\boldsymbol{x}^{e} + \mathbf{H}^{c}_{F_{FA}}\mathbf{F}^{c,k}_{FA} + \mathbf{H}_u\mathbf{u}^k\\ \frac{d\mathbf{F}^{e}_{Elas}}{dt} &\simeq \mathbf{R}^{e}_{x}\boldsymbol{x}^{e} + \mathbf{R}^{e}_{F_{Elas}}\mathbf{F}^{e}_{Elas}\end{aligned}\right\} set\ 2 \quad (5)$$

The first set of differential equations are a re-representation of the original state equations in equation (1) and (2) which are apparently linear in terms of the auxiliary variables and input. Here, $\boldsymbol{x}^{c,k}=\left(\boldsymbol{x}^{c,k\mathbf{T}}_1 \ \cdots \ \boldsymbol{x}^{c,k\mathbf{T}}_{N_c}\right)^{\mathbf{T}}\in\Re^{3N_c\times 1}$ is a vector containing the 3-D coordinates of all the cell membrane nodes. $\mathbf{F}^{c,k}_{Cort-Elas}\in\Re^{3N_C\times 1}$ is a vector comprising cortical tension and elastic energy forces for all the cell nodes ($i=1,\cdots,N_C$), $\mathbf{F}^{c,k}_{FA}\in\Re^{3N_C\times 1}$ is a vector of focal adhesion forces at all the cell nodes, $\mathbf{u}^k$ is an input vector containing all the lamellipodium forces ($\mathbf{F}^{c,k}_{L,i}$), and $\mathbf{W}^{c}_{CE},\mathbf{W}^{c}_{FA}$ and $\mathbf{L}_\mathbf{C}$ are constant matrices of consistent dimensions. $\boldsymbol{x}^{e}=\left(\boldsymbol{x}^{e\mathbf{T}}_1 \ \cdots \ \boldsymbol{x}^{e\ \mathbf{T}}_{N_e}\right)^{\mathbf{T}}\in\Re^{3N_e\times 1}$ is a vector containing the 3-D coordinates of all the ECM nodes. And $\mathbf{F}^{e}_{Elas}\in\Re^{3N_e\times 1},\mathbf{F}^{e}_{FA}\in\Re^{3N_e\times 1}$ are vectors consisting of all the elastic energy forces and focal adhesion forces acting on the ECM network, respectively.

The second set of differential equations represent the transition of auxiliary state variables estimated through linear regressions. Here, $\mathbf{R}^{*}_{*}\in\Re^{3N_e\times 3N_e}$, $\mathbf{Q}^{*}_{*}\in\Re^{3N_c\times 3N_c}$, $\mathbf{H}^{*}_{*}\in\Re^{3N_c\times 3N_c}$ (*-corresponding various subscripts and superscripts) are high-dimensional parameter matrices. These matrices can be determined from numerical data created by simulation of the original dynamical system.

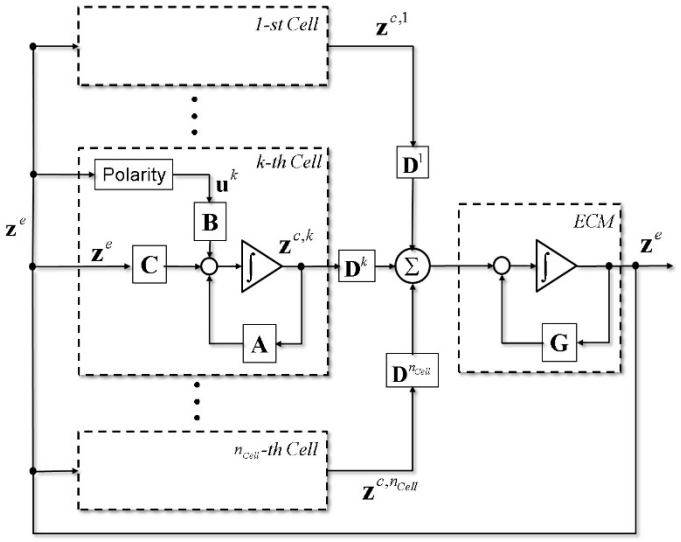

FIGURE 2: BLOCK DIAGRAM OF LATENT VARIABLE SUPERPOSITION MODEL REPRINTED FROM (16)

However, due to the high dimension, these parameters matrices could be difficult to determine explicitly. However if the system is represented in a lower dimensional space, the high dimensional regression coefficient matrices need not be computed explicitly. This is discussed further in the subsequent section.

Both differential equation sets presented in equations (4) and (5) are linear representing different (or dual) facets of the original nonlinear system viewed from the augmented space and providing a richer representation of the nonlinearity [16].

LATENT VARIABLE TRANSFORMATION

A drawback of the Dual Faceted linearization method is the increase of variables. Auxiliary variables are not independent variables, and the representation with two sets of differential equations is highly redundant in a sense. Some variables may be collinear, and the differential equations may contain similar modes. These similar modes and collinear variables can be eliminated by using latent variable models [17,18].

From the simulation of the state equations (1) and (2), we can obtain sample data of the state variables in $\boldsymbol{x}^{c,k}$ and $\boldsymbol{x}^{e}$. We can also calculate auxiliary variables in $\mathbf{F}^{c,k}_{Cort-Elas}$, $\mathbf{F}^{c,k}_{FA}$, $\mathbf{F}^{e}_{Elas}$ using the nonlinear relationships with the ECM and membrane coordinates [14,15]. The training data consists of *T~3600* time sample points from state and auxiliary variables of a *K=1-2* cells embedded in an ECM environment. The simulation is repeated N ~ 10 times with the cell(s) embedded in distinct locations within the ECM.

Let $\boldsymbol{\zeta}^{\boldsymbol{c.k}}$ be the augmented variable vector containing membrane node coordinates and forces of the *k*-th cell.

$$\zeta^{c,k} = \begin{pmatrix} \boldsymbol{x}^{c,k} \\ \mathbf{F}^{c,k}_{Cort-Elas} \\ \mathbf{F}^{c,k}_{FA} \end{pmatrix} \in \Re^{9N_c \times 1} \tag{6}$$

Here vector $\mathbf{u^k}$ (containing the cell's lamellipodial forces) is treated as an input variable that is excluded from the augmented state space.

Similarly let $\boldsymbol{\zeta^e}$ be the augmented variable vector containing ECM node coordinates and forces:

$$\zeta^{e} = \begin{pmatrix} \boldsymbol{x}^{e} \\ \mathbf{F}^{e}_{Elas} \end{pmatrix} \in \Re^{6N_e \times 1} \tag{7}$$

Focal adhesion forces contained within vector $\mathbf{F}^e_{FA}$ can be mapped to the corresponding focal adhesion forces at each cell and, thereby, excluded from the augmented space of the ECM. Let the covariance matrices $\mathbf{C}^c_{\zeta\zeta}, \mathbf{C}^e_{\zeta\zeta}$ be:

$$\begin{aligned} C^c_{\zeta\zeta} &= \frac{1}{K \cdot N \cdot T} \sum_{k=1}^{K} \sum_{n=1}^{N} \sum_{t=1}^{T} \tilde{\zeta}^{c,k,n}(t) \tilde{\zeta}^{c,k,n}(t)^{\mathbf{T}} \\ C^e_{\zeta\zeta} &= \frac{1}{N \cdot T} \sum_{n=1}^{N} \sum_{t=1}^{T} \tilde{\zeta}^{e,n}(t) \tilde{\zeta}^{e,n}(t)^{\mathbf{T}} \end{aligned} \tag{8}$$

Here $\tilde{\zeta}^{c,n,k}(t)$ represents the mean centered t-th time sample (of augmented variable vector $\boldsymbol{\zeta}^{c.k}$) for the k-th cell in the n-th simulation and $\tilde{\zeta}^{e,n}(t)$ represents the mean centered t-th time sample (of the augmented variable vector $\boldsymbol{\zeta^e}$) in the n-th simulation.

We transform the augmented linearized system one in the latent variable space spanned by eigenvectors $\mathbf{V}^c = \begin{pmatrix} \mathbf{V}^{c\mathbf{T}}_x & \mathbf{V}^{c\ \mathbf{T}}_{F_{CE}} & \mathbf{V}^{c\ \mathbf{T}}_{F_{FA}} \end{pmatrix}^{\mathbf{T}} \in \Re^{m_c \times m_c}$ and $\mathbf{V}^e = \begin{pmatrix} \mathbf{V}^{e\mathbf{T}}_x & \mathbf{V}^{e\ \mathbf{T}}_{F_{Elas}} \end{pmatrix}^{\mathbf{T}} \in \Re^{m_e \times m_e}$ of the covariance matrices $\mathbf{C}^c_{\zeta\zeta}, \mathbf{C}^e_{\zeta\zeta}$ respectively:

$$\mathbf{z}^{c,k} = \begin{pmatrix} \mathbf{V}^c_x \\ \mathbf{V}^c_{F_{CE}} \\ \mathbf{V}^c_{F_{FA}} \end{pmatrix}^{\mathbf{T}} \begin{pmatrix} \boldsymbol{x}^{c,k} \\ \mathbf{F}^{c,k}_{Cort-Elas} \\ \mathbf{F}^{c,k}_{FA} \end{pmatrix} \in \Re^{m_c \times 1}, \mathbf{z}^{e} = \begin{pmatrix} \mathbf{V}^e_x \\ \mathbf{V}^e_{F_{Elas}} \end{pmatrix}^{\mathbf{T}} \begin{pmatrix} \boldsymbol{x}^{e} \\ \mathbf{F}^{e}_{Elas} \end{pmatrix} \in \Re^{m_e \times 1} \tag{9}$$

Here, $m_c << 3N_c$ and $m_e << 3N_e$. In addition, the original data of $\boldsymbol{\zeta^{c.k}}$ and $\boldsymbol{\zeta^e}$ can be approximated with latent variables:

$$\zeta^{c,k} = \begin{pmatrix} \mathbf{V}^c_x \\ \mathbf{V}^c_{F_{CE}} \\ \mathbf{V}^c_{F_{FA}} \end{pmatrix} \mathbf{z}^{c,k}, \ \zeta^{e} = \begin{pmatrix} \mathbf{V}^e_x \\ \mathbf{V}^e_{F_{Elas}} \end{pmatrix} \mathbf{z}^{e} \tag{10}$$

Differentiating the latent state vector $\mathbf{z}^{c.k}$ and substituting equations (4), (5) and (10) yields:

$$\begin{aligned} \frac{d\mathbf{z}^{c,k}}{dt} &= \mathbf{V}^{c\mathbf{T}}_x \frac{d\boldsymbol{x}^{c,k}}{dt} + \mathbf{V}^{c\ \mathbf{T}}_{F_{CE}} \frac{d\mathbf{F}^{c,k}_{Cort-Elas}}{dt} + \mathbf{V}^{c\ \mathbf{T}}_{F_{FA}} \frac{d\mathbf{F}^{c,k}_{FA}}{dt} \\ &= \mathbf{A}\,\mathbf{z}^{c,k} + \mathbf{B}\,\mathbf{u}^k + \mathbf{C}\mathbf{z}^e \end{aligned} \tag{11}$$

Where:

$$\begin{aligned} \mathbf{A} &= \mathbf{V}^{c\mathbf{T}}_x (\mathbf{W}^c_{CE} \mathbf{V}^c_{F_{CE}} + \mathbf{W}^c_{FA} \mathbf{V}^c_{F_{FA}}) \\ &\quad + \mathbf{V}^{c\ \mathbf{T}}_{F_{CE}} (\mathbf{Q}^c_x \mathbf{V}^c_x + \mathbf{Q}^c_{F_{CE}} \mathbf{V}^c_{F_{CE}}) + \mathbf{V}^{c\ \mathbf{T}}_{F_{FA}} (\mathbf{H}^c_x \mathbf{V}^c_x + \mathbf{H}^c_{F_{FA}} \mathbf{V}^c_{F_{FA}}) \\ \mathbf{B} &= \mathbf{V}^{c\mathbf{T}}_x \mathbf{L}_c + \mathbf{V}^{c\ \mathbf{T}}_{F_{CE}} \mathbf{Q}_u + \mathbf{V}^{c\ \mathbf{T}}_{F_{FA}} \mathbf{H}_u \\ \mathbf{C} &= \mathbf{V}^{c\ \mathbf{T}}_{F_{FA}} \mathbf{H}^e_x \mathbf{V}^e_x \\ \mathbf{G} &= \mathbf{V}^{e\mathbf{T}}_x \mathbf{W}^e_{Elas} \mathbf{V}^e_{F_{Elas}} + \mathbf{V}^{e\ \mathbf{T}}_{F_{Elas}} (\mathbf{R}^e_x \mathbf{V}^e_x + \mathbf{R}^e_{F_{Elas}} \mathbf{V}^e_{F_{Elas}}) \\ \mathbf{D}^k &= \mathbf{V}^{e\mathbf{T}}_x \mathbf{W}^e_{FA} \mathbf{P}^k_{map} \mathbf{V}^c_{F_{FA}} \end{aligned} \tag{12}$$

Here, $\mathbf{P}^k_{map} \in \Re^{3N_e \times 3N_c}$ is a parameter matrix (consisting of either 0 or -1 elements) which maps the membrane focal adhesion forces of the k-th cell ($\mathbf{F}^{c,k}_{FA}$) to the corresponding ECM focal adhesion forces ($\mathbf{F}^e_{FA}$) as discussed in reference [16].

Differentiating the latent state vector $\mathbf{z}^e$ and substituting equation (4), (5) and (10) yields:

$$\frac{d\mathbf{z}^e}{dt} = \mathbf{V}^{e\mathbf{T}}_x \frac{d\boldsymbol{x}^e}{dt} + \mathbf{V}^{e\ \mathbf{T}}_{F_{Elas}} \frac{d\mathbf{F}^e_{Elas}}{dt} = \mathbf{G}\,\mathbf{z}^e + \sum_{k=1}^{n_{cell}} \mathbf{D}^k \mathbf{z}^{c,k} \tag{13}$$

Equations (11) and (13) provide an accurate representation of the nonlinear ECM dynamics reduced to a compact, low-dimensional model. This model provides not only a low-dimensional modular structure for efficient computation, but also contains natural insights into the interactions among the multiple cells. Figure 2 shows the dynamic interactions in block diagram form based on equations (11) and (13) and represents our mathematical structure of modular components whose simplified interactions still reflect the physical mechanisms within the system. As can be seen, the actions taken by all the cells are integrated into the global ECM state transition, which is fed back to the individual cells. Therefore, each cell is connected to other cells through the global feedback of the ECM latent state $\mathbf{z^e}$.

The input $\mathbf{u^k}$ pertains to the lamellipodium forces, which are generated in a particular side or direction of the cell. Among others, the polarity of a cell is important to determine on which side of the cell lamellipodia are formed. It has been reported that the local stiffness of ECM is a major factor that determines the cell polarity [19]. The functional relation between lamellipodium force generation $\mathbf{u^k}$ and properties and state of the ECM stress-strain field $\mathbf{z^e}$ are described in [16]. This can be interpreted as a feedback controller residing in each cell model.

LEAST SQUARES ESTIMATION FOR IDENTIFICATION OF THE PARAMETER MATRICES INVOLVED IN THE LATENT SPACE STATE EQUATIONS

Since the system is represented in a lower dimensional space, the high dimensional regression coefficient matrices ($\mathbf{R}_*^*, \mathbf{Q}_*^*, \mathbf{H}_*^*$) are not computed explicitly. Instead, the lower dimension coefficient matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{G}$ are computed directly from numerical simulation data transformed into the latent variable space. Recall that training data consists of *T~3600* time sample points from state and auxiliary variables of a *K=1-2* cells embedded in an ECM environment. The simulation is repeated N ~ 10 times with the cell(s) embedded in distinct locations within the ECM.

We define transformed data set:

$$Z_{Tr} = \left\{ \left(\mathbf{z}^{e,n}(t), \mathbf{z}^{c,k,n}(t), \mathbf{u}^{k,n}(t), d\mathbf{z}^{e,n}/dt, d\mathbf{z}^{c,k,n}/dt \right) \middle| k = 1,\cdots,K, n = 1,\cdots,N, t = 1,\cdots,T \right\} \quad (14)$$

Here superscripts *k, n* signify the *k*-th cell within the *n*-th simulation. We combine parameter matrices from equation (11) $\mathbf{M} \triangleq [\mathbf{A} \quad \mathbf{B} \quad \mathbf{C}] \in \Re^{m_c \times (m_c + N_c + m_e)}$ and variables into $\xi^{k,n}(t) = \left(\mathbf{z}^{c,k,n}(t)^{\mathrm{T}} \quad \mathbf{u}^{k,n}(t)^{\mathrm{T}} \quad \mathbf{z}^{e,n}(t)^{\mathrm{T}} \right)^{\mathrm{T}} \in \Re^{(m_c + N_c + m_e) \times 1}$. The parameter matrix $\mathbf{M}$ can be optimized so that the mean squared error of predicting $d\mathbf{z}^{c,k,n}/dt$ may be minimized:

$$\mathbf{M}^0 = \arg\min_{\mathbf{M}} \frac{1}{K \cdot N \cdot T} \sum_{k=1}^{K} \sum_{n=1}^{N} \sum_{t=1}^{T} \left\| \left. \frac{d\mathbf{z}^{c,k,n}}{dt} \right|_t - \mathbf{M} \xi^{k,n}(t) \right\|^2 \quad (15)$$

Using the standard least squared estimation and assuming that the sample data sufficiently spans the dimension of vector $\xi^{k,n}(t)$, we can obtain:

$$\mathbf{M}^0 = \left(\sum_{k=1}^{K} \sum_{n=1}^{N} \sum_{t=1}^{T} \left. \frac{d\mathbf{z}^{c,k,n}}{dt} \right|_t \xi^{k,n}(t)^{\mathrm{T}} \right) \left(\sum_{k=1}^{K} \sum_{n=1}^{N} \sum_{t=1}^{T} \xi^{k,n}(t) \xi^{k,n}(t)^{\mathrm{T}} \right)^{-1} \quad (16)$$

Similarly least squares estimate matrix $\mathbf{G}$ from equation (13) is given by:

$$\mathbf{G}^0 = \left(\sum_{n=1}^{N} \sum_{t=1}^{T} \boldsymbol{\delta}^n(t) \mathbf{z}^{e,n}(t)^{\mathrm{T}} \right) \left(\sum_{n=1}^{N} \sum_{t=1}^{T} \mathbf{z}^{e,n}(t) \mathbf{z}^{e,n}(t)^{\mathrm{T}} \right)^{-1} \quad (17)$$

Where $\boldsymbol{\delta}^n(t) = \left. d\mathbf{z}^{c,k,n}/dt \right|_t - \sum_{k=1}^{K} \mathbf{D}^k \mathbf{z}^{c,k,n}$ and $\mathbf{D}^k$'s are known matrices as defined in (12).

In summary, in order to to compute coefficient matrices **A**, **B**, **C**, **G** we do the following:

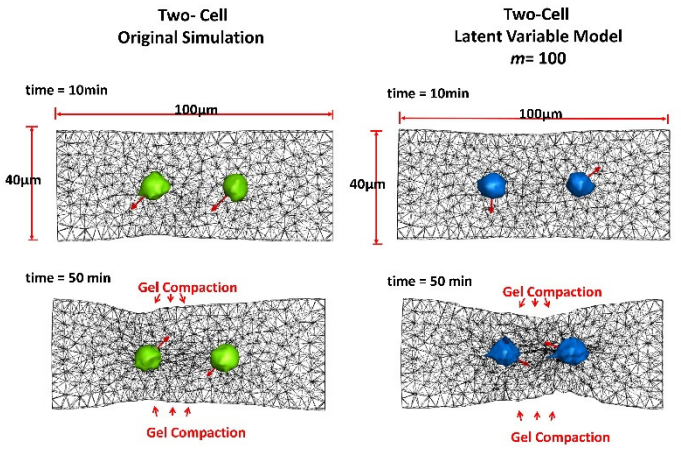

FIGURE 3: COMPARISON OF ECM COMPACTION FOR NONLINEAR COMPUTATIONAL MODEL AND LINEAR LATENT VARIABLE MODEL

	20 LV model	**100LV model**	**Real Simulation**
Computation time	2 min	20 min	10 hours

TABLE 1: SUMMARY OF THE COMPUTATION TIME OF THE LATENT VARIABLE MODEL COMPARED TO THE ORIGINAL FULL-SCALE SIMULATION

1) Create data by simulating the original state equations, equations (1) and equations (2), using the full-scale, nonlinear model, as described previously.
2) Compute covariance matrices $C_{\zeta\zeta}^c$, $C_{\zeta\zeta}^e$, and obtain eigenvalues and eigenvectors $\mathbf{V}^c$ and $\mathbf{V}^c$, as described previously.
3) Transform the data of the augmented state variables to latent space ($\mathbf{z}^{c,k,n}(t)$ and $\mathbf{z}^{e,n}(t)$) using the orthogonal matrices $\mathbf{V}^c$ and $\mathbf{V}^c$.
4) Compute time derivatives $d\mathbf{z}^{c,k,n}/dt$ and $d\mathbf{z}^{e,n}/dt$, using latent space time samples and form a dataset: $Z_{Tr} = \left\{ \left(\mathbf{z}^{e,n}(t), \mathbf{z}^{c,k,n}(t), \mathbf{u}^{k,n}(t), d\mathbf{z}^{e,n}/dt, d\mathbf{z}^{c,k,n}/dt \right) \middle| k = 1,\cdots,K, n = 1,\cdots,N, t = 1,\cdots,T \right\}$. Using Z_{Tr}, identify the parameter matrices **A**, **B**, **C**, **G** involved in the latent space state equations using Least Squares Estimate.

Parameter matrices $\mathbf{A} \in \Re^{m_c \times m_c}, \mathbf{B} \in \Re^{m_c \times 3N_c}, \mathbf{C} \in \Re^{m_c \times m_e}, \mathbf{G} \in \Re^{m_e \times m_e}$ are much lower in dimension than the regression coefficient matrices $\mathbf{R}_*^*, \mathbf{Q}_*^*, \mathbf{H}_*^*$ given in equations (5). Therefore, fewer data points allow us to determine these parameter matrices in the latent space. It should be noted that matrices $\mathbf{D}^k$'s are of high dimension, but are not computed with regression since they consist of known matrices as shown in equation (12).

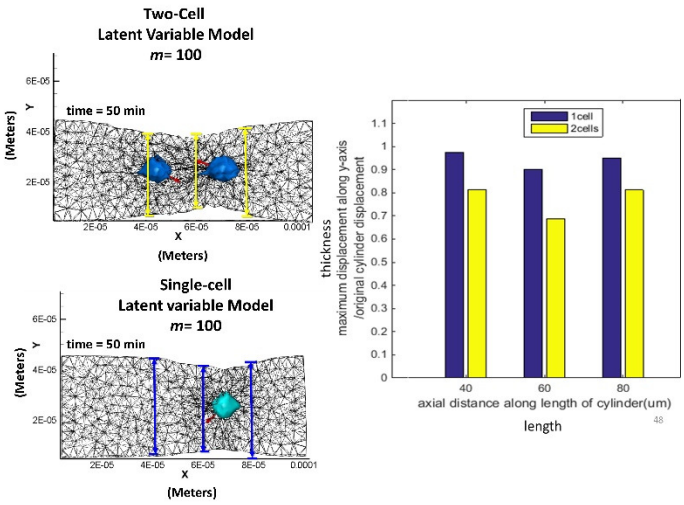

FIGURE 4: COMPARISON OF ECM CONTRACTION AT DIFFERENT POINTS ALONG THE LONGITUDINAL AXIS

RESULTS

Figure 3 shows simulation experiments of two cells interacting with 3D ECM. Two cells, initially with an initial spherical shape, are embedded in the ECM fiber network of a cylindrical space that measures 40 μm in diameter and 100 μm in length. Cells began to interact with the ECM fibers immediately, formed focal adhesions, and deformed both their own shapes and the ECM fiber network. Polarity directions of both cells (red arrows initially pointing in arbitrary directions) shift to point inward, indicating that larger stresses are detected in the area between the cells .The left hand side plots with green cells were generated with the original full scale computation using (1) and (2) while the right hand side plots with blue cells were the simulation using the dual faceted linearization and latent variable modeling based on (11) and (13). The low-order latent model could successfully reproduce the ground-truth, full-scale simulation results in a significantly less amount of time. Table 1 give a summary of the computation time of the latent variable model using $m= m_c + m_e = 20$ latent variables and m= 100 latent variables compared to the original full-scale simulation.

Figure 4 shows the ECM gel compaction by a single cell alone compared with two cells. The compaction by a single cell is significantly smaller as quantified by the thickness comparison of the ECM along the longitudinal (X) axis. The maximum contraction for 1 cell along the axis is maximum contraction 90% which the maximum contraction 67% . It should be noted that the compaction by the two cells is more than twice larger than that of the single cell. This implies that the two cells amplified their reactions to the ECM compaction with each other, exhibiting a collective behavior.

CONCLUSION

The collective ECM compaction by multiple cells was predicted through superposition of individual cells contributions. This was made possible with the Dual Faceted Linearization. As applied to the analysis of multi-cell ECM compaction, linear augmented equations describing single cell-ECM interactions were derived from DF linearization, and then converted to a reduced-order linear representation by transformation onto a basis of eigenvectors derived from simulated data set. DF Linearization allows for the evolution of independent and auxiliary states to be described within a lower dimensional linear manifold. The resulting reduced order latent model is capable of reproducing nonlinear dynamics, and the linearized structure of individual models facilitated their integration to describe cell behaviors. The prediction of collective behaviors was achieved by superposing contributions of individual cells represented by latent variables $\mathbf{z}^{c,k}$, which evolves based on their own dynamics in response to the global ECM state represented by latent variable $\mathbf{z}^{e}$. The presented method for predicting collective behaviors of cell-mediated ECM gel compaction is scalable. Since the individual cell-ECM interactions are local computations the computational complexity does not increase exponentially, although the number of cells increases. Current and future work includes increasing the number of cells within the latent variable simulation [16]. Since computing a ground truth simulation with a large number of cells is impractical and sometimes infeasible, the proposed work can be used to predict large scale multi-cell interactions within reasonable computational time frame and can be compared with experimental results.

ACKNOWLEDGMENTS

The authors would like to thank Prof. Roger D. Kamm (MIT) and Taher Saif (UIUC) for their biological insights and advice of the studied system. Funding: The authors acknowledge support from the National Science Foundation (NSF), Science and Technology Center (STC) on Emergent Behaviors in Integrated Cellular Systems (EBICS) under Grant CBET-0939511.

REFERENCES

[1] Doljanski, F., 2004, "The Sculpturing Role of Fibroblast-like Cells in Morphogenesis," Perspect. Biol. Med., **47**(3), pp. 339–356.

[2] Davis, G. E., and Camarillo, C. W., 1995, "Regulation of Endothelial Cell Morphogenesis by Integrins, Mechanical Forces, and Matrix Guidance Pathways," Exp. Cell Res., **216**(1), pp. 113–123.

[3] Li, B., and Wang, J. H.-C., 2011, "Fibroblasts and Myofibroblasts in Wound Healing: Force Generation and Measurement," J Tissue Viability, **20**(4), pp. 108–120.

[4] Grinnell, F., 1994, "Fibroblasts, Myofibroblasts, and Wound Contraction," J. Cell Biol., **124**(4), pp. 401–404.

[5] Enever, P. A. J., Shreiber, D. I., and Tranquillo, R. T., 2002, "A Novel Implantable Collagen Gel Assay for Fibroblast Traction and Proliferation during Wound Healing," J. Surg. Res., **105**(2), pp. 160–172.

[6] Ma, X., Schickel, M. E., Stevenson, M. D., Sarang-Sieminski, A. L., Gooch, K. J., Ghadiali, S. N., and Hart, R. T., 2013, "Fibers in the Extracellular Matrix Enable Long-Range Stress Transmission between Cells," Biophys J, **104**(7), pp. 1410–1418.

[7] Stevenson, M. D., Sieminski, A. L., McLeod, C. M., Byfield, F. J., Barocas, V. H., and Gooch, K. J., 2010, "Pericellular Conditions Regulate Extent of Cell-Mediated Compaction of Collagen Gels," Biophys. J., **99**(1), pp. 19–28.

[8] Aghvami, M., Barocas, V. H., and Sander, E. A., 2013, "Multiscale Mechanical Simulations of Cell Compacted Collagen Gels," J Biomech Eng, **135**(7), pp. 0710041–0710049.

[9] Reinhardt, J. W., Krakauer, D. A., and Gooch, K. J., 2013, "Complex Matrix Remodeling and Durotaxis Can Emerge from Simple Rules for Cell-Matrix Interaction in Agent-Based Models," Journal of biomechanical engineering, **135**(7), p. 071003.

[10] Reinhardt, J. W., and Gooch, K. J., 2014, "Agent-Based Modeling Traction Force Mediated Compaction of Cell-Populated Collagen Gels Using Physically Realistic Fibril Mechanics," J Biomech Eng, **136**(2), p. 021024.

[11] Moon, A. G., and Tranquillo, R. T., 1993, "Fibroblast-Populated Collagen Microsphere Assay of Cell Traction Force: Part 1. Continuum Model," AIChE J., **39**(1), pp. 163–177.

[12] Barocas, V. H., and Tranquillo, R. T., 1997, "An Anisotropic Biphasic Theory of Tissue-Equivalent Mechanics: The Interplay among Cell Traction, Fibrillar Network Deformation, Fibril Alignment, and Cell Contact Guidance," Journal of biomechanical engineering, **119**(2), pp. 137–145.

[13] Tranquillo, R. T., Durrani, M. A., and Moon, A. G., 1992, "Tissue Engineering Science: Consequences of Cell Traction Force," Cytotechnology, **10**(3), pp. 225–250.

[14] Kim, M.-C., Whisler, J., Silberberg, Y. R., Kamm, R. D., and Asada, H. H., 2015, "Cell Invasion Dynamics into Three Dimensional Extracellular Matrix Fiber Network(Accepted 9/2015)," PLoS computational biology.

[15] Kim, M.-C., Silberberg, Y. R., Abeyaratne, R., Kamm, R. D., and Asada, H. H., 2018, "Computational Modeling of Three-Dimensional ECM-Rigidity Sensing to Guide Directed Cell Migration," Proceedings of the National Academy of Sciences, p. 201717230.

[16] Mayalu, M. N., Asada, H. H., and Kim, M.-C., "Multi-Cell ECM Compaction is Predictable via Superposition of Nonlinear Cell Dynamics Linearized in Augmented State Space," Cell Systems (Under Review).

[17] Kruger, U., and Xie, L., *Statistical Monitoring of Complex Multivariate Processes: With Applications in Industrial Process Control*.

[18] Kerschen, G., Golinval, J., Vakakis, A. F., and Bergman, L. A., 2005, "The Method of Proper Orthogonal Decomposition for Dynamical Characterization and Order Reduction of Mechanical Systems: An Overview," Nonlinear Dyn, **41**(1–3), pp. 147–169.

[19] Borau, C., Kamm, R. D., and García-Aznar, J. M., 2011, "Mechano-Sensing and Cell Migration: A 3D Model Approach," Phys Biol, **8**(6), p. 066008.