Reconstruction of the phase dynamics of the somitogenesis clock oscillator

Lucas J. Morales Moya
J. Kim Dale
Philip J. Murray

August 21, 2019

Abstract

In this study we develop a computational framework for the reconstruction of phase dynamics in spatio-temporally oscillating systems and use it to study the dynamics of the somitogenesis clock oscillator. Our understanding of the somitogenesis clock, a developmental oscillator found in the vertebrate embryo, has been revolutionised by the development of real time reporters of clock gene expression. However, the signals obtained from the real time reporters are typically noisy, nonstationary and spatiotemporally dynamic and there are open questions with regard to how post-processing can be used to both improve the insight gained from a given experiment and to constrain theoretical models. In this study we present a methodology, which is a variant of empirical mode decomposition, that reconstructs the phase dynamics of the somitogenesis clock. After validating the methodology using synthetic datasets, we define a set of metrics that use the reconstructed phase profiles to infer biologically meaningful quantities. We perform experiments that measure signal from a real time reporter of the somitogenesis clock and reconstruct the phase dynamics. Application of the defined metrics yields results that are consistent with previous experimental observations. Moreover, we extend previous work by developing a gradient descent method for defining unbiased kymographs and showing that boundary conditions are non-homogeneous. By studying phase dynamics along phase gradient descent trajectories, we show that, consistent with a previous theoretical model, the oscillation frequency is inversely correlated with the phase gradient but that the coefficient is not constant in time. The proposed methodology provides a tool kit for that can be used in the analysis of future experiments and the quantitative observations can be used to guide the development of future mathematical models.

1 Introduction

During development of the vertebrate embryo, the head-tail axis sequentially segments into pairs of segments at regular intervals in time. Underlying tem-

poral periodicity is a molecular oscillator known as the segmentation clock that is characterised by waves of gene expression (Palmeirim et al., 1997; Lauschke et al., 2013; Soroldoni et al., 2014) that are observed in presomitic mesoderm (PSM) and come to rest at the boundary of future segment.

The development of new PSM culture systems (Lauschke et al., 2013; Tsiairis and Aulehla, 2016, e.g.) has allowed much higher throughput analyses of spatiotemporal dynamics than was previously possible. Of particular relevance to this study is an *ex vivo* mPSM explant culture system in which a small, dissected part of the PSM from a live reporter mouse is cultured as a monolayer. When cultured under appropriate conditions, the mPSM explants exhibit spatiotemporal oscillations of gene expression. An open question in the field is how the vast amounts of real-time data can be processed and quantified so as to maximise the inference that can be made from real-time reporter studies.

The development of real-time reporters of clock gene expression in multiple vertebrate species has revolutionised the field of somitogenesis research (Soroldoni and Oates, 2011; Aulehla et al., 2008; Masamizu et al., 2006). Numerous methods have been used to process real time reporter signal so as to yield biologically meaningful inference: the time that elapses between peaks of gene expression (Lauschke et al., 2013; Webb et al., 2016; Hubaud et al., 2017); the Fourier transform (Tsiairis and Aulehla, 2016); the Wavelet Transform (Soroldoni et al., 2014; Webb et al., 2016; Hubaud et al., 2017); the Hilbert Transform (Lauschke et al., 2013); moving averages (Sonnen et al., 2018); moving average and a Savitzky-Golay filter (Matsumiya et al., 2018); representation as a harmonic oscillator (Delaune et al., 2012).

The computation of the phase of an oscillating signal allows for the relative progression through an oscillatory cycle to be quantified. In complex spatiotemporal systems, the phase dynamics of the signal can allow the identification of new phenomena without the need to necessarily fully understand the mechanisms (e.g. genetic components) underpinning the signal. Moreover, the inference of phase dynamics from an oscillatory signal provides a means to link experiment observations with theory.

There are numerous well-established techniques that allow the representation of temporally oscillating signals. The Fourier transform decomposes a signal into a linear sum of harmonic functions. However, for a non-stationary signal, such as a chirp signal where the frequency varies linearly time, the Fourier spectrum does not provide an intuitive representation of signal dynamics and is not particularly useful in the recovery of phase dynamics. The Wavelet Transform (Mallat, 1999), which represents a signal as a sum of predefined wavelet functions, allows for a localised description of a non-stationary signals. The results of a wavelet analysis are typically presented via a frequency spectrogram. To infer phase dynamics of the underlying signal, methods such as ridge detection are applied (e.g. Harang et al., 2012).

The application of the Hilbert transform to a signal is an alternative to templated-based methods such as the Fourier and Wavelet Transforms. Here, a

periodic signal is represented by

$$s(t) = r(t)\cos\left(\theta(t)\right), \tag{1}$$

where $r(t)$ is the instantaneous amplitude and $\theta(t)$ the instantaneous phase. Such a signal is considered to be a monochromatic signal, as it only contains a single frequency at any time point. Whilst the Hilbert transform provide provides a natural definition of oscillator phase, its application to real-world signals is complicated by the presence of trends, noise and nonstationarity (Pikovsky et al., 2001; Huang et al., 1998).

To ensure that a given signal can be studied systematically using the Hilbert Transform, Huang et al. (1998) introduced a data-driven, adaptive signal analysis technique known as Empirical Mode Decomposition (EMD). This technique, which does not require basis functions to be predefined, decomposes a signal into a set of intrinsic mode functions (IMF) that represent oscillations on distinct timescales. Each of the IMFs, which are monochromatic and can have meaningful physical/biological interpretations (Huang et al., 1998), can subsequently be analysed using the Hilbert Transform. Important limitations of EMD are its sensitivity to noise and mode mixing (Wu et al., 2009; Rehman et al., 2013). These limitations hinder the application of EMD in many practical situations.

To address the limitations of EMD, the original formulation has been modified in numerous ways (Rilling et al., 2007; Rehman and Mandic, 2010, 2009; Nunes et al., 2003; Bhuiyan et al., 2008; Wu et al., 2009; Riffi et al., 2015; Wu and Huang, 2009; Rehman et al., 2013). Multivariate empirical mode decomposition (MEMD) is a modification of EMD in which a multi-variate signal can be represented as sum of intrinsic mode functions (Rehman and Mandic, 2009). Noise assisted multivariate empirical mode decomposition (NA-MEMD) is a modification of EMD in which a univariate signal is processed by introducing Gaussian noise in neighbouring channels and then applying MEMD such that the relative position of signal in the different modes is regularised. NA-MEMD is well suited to the task of decomposing noisy time series originating from nonlinear, non-stationary oscillators.

EMD has been presented as an alternative to methods such as Wavelet Transform (Huang et al., 1998). Nonetheless, the Wavelet Transform has been widely used to study spatiotemporal dynamics (Kikuchi and Wang, 2010). The Wavelet Transform allows the study of the frequencies that compose the signal at a given point. However, the application of this transform increases the dimensionality of the analysis, which requires the development of sophisticated metrics. Additionally, he dependence on the wavelet of choice, which can left out relevant information of nonlinear signals, and the indirect measurement of the instantaneous phase are the main drawbacks for such methods (Huang et al., 1998; Alegre-Cortés et al., 2016; Hirsh et al., 2018).

EMD has been extended to the study of spatio-temporal signals. Bi-dimensional and tri-dimensional EMD have been developed to represent a spatiotemporal wave as a sum of spatiotemporal intrinsic mode functions with the objective being to decompose a signal into characteristic length-scales (Wu and Huang,

2009; Schmitt et al., 2014; He et al., 2017). Multidimensional Empirical Mode Decomposition (Wu et al., 2009) has been developed, in the context of image analysis, as a generalisation of EMD to two and three spatial dimensions (Nunes et al., 2003; Chen and Jeng, 2014). EMD has been combined with Principal Component Analysis or Empirical Orthogonal Functions to account for spatial structure (Park et al., 2013; Davies and James, 2014; Wu et al., 2016). Spatiotemporal MEMD is a multivariate EMD that has been used to incorporate spatial information in Brain Computer Interfaces (Park et al., 2013; Davies and James, 2014).

In this study we develop a methodology for the reconstruction of phase dynamics in an experimentally-motivated situation where a spatio-temporally oscillating signal is modulated by a time-dependent envelope. The general problem is to reconstruct the phase, $\theta(x,t)$, from a signal of the form

$$s(\underline{x},t) = r(\underline{x},t)\cos(\theta(\underline{x},t)) + \xi_\sigma(\underline{x},t), \tag{2}$$

where

$$\theta(\underline{x},t) = \int_0^t \omega(\underline{x},t)dt, \quad \omega(\underline{x},t) > 0, \tag{3}$$

where $r(\underline{x},t)$ represents the signal amplitude, $\theta(\underline{x},t)$ represents the phase, $\omega(\underline{x},t)$ represents the frequency and $\xi_\sigma(\underline{x},t)$ represents Gaussian noise of strength σ. The layout of the paper is as follows: in Section 2 we outline methods; in Section 3 develop and validate a methodology for phase reconstruction that is based on EMD and apply it to experimental data from a real time reporter of the somitogenesis clock; and, finally, in Section 4 we conclude with a discussion.

2 Methods

2.1 A synthetic dataset

2.1.1 Model equations

We generate a synthetic dataset that exhibits the major features of real time reports of the somitogenesis clock in mPSM explants: an amplitude gradient, a frequency gradient, initially spatially homogeneous oscillations followed by the progressive loss of signal at the boundary.

Let the instantaneous phase, $\theta(\underline{x},t)$, be given by

$$\theta(\underline{x},t_{seg}) = \omega t, \quad \forall \;\; 0 < t < t_{seg}, \quad \underline{x} \in D. \tag{4}$$

where t_{seg} represent the time at which segmentation begins. Suppose that $\theta(\underline{x},t)$ satisfies the partial differential equation

$$\frac{\partial\theta}{\partial t} = \omega + A\nabla^2\theta + B\nabla\theta\cdot\nabla\theta, \quad t > t_{seg}, \quad \underline{x} \in D, \tag{5}$$

where ω is the natural oscillation frequency, and A and B are coupling parameters (Murray et al., 2011). The domain D is defined to be the disk of radius R, i.e.

$$D = \{\underline{x} \in \mathbb{R}^2 : |\underline{x}| < R\}. \tag{6}$$

The boundary conditions are given by

$$\theta(|\underline{x}| = R, t) = \omega t_{seg}, \quad \forall t > t_{seg}, \tag{7}$$

and the initial conditions are

$$\theta(\underline{x}, t_{seg}) = \omega t_{seg}, \quad \forall \underline{x} \in D. \tag{8}$$

To imitate the signal decay after arrest, the amplitude is defined to be

$$r(\underline{x}, t) = \begin{cases} 1 + k_1 \frac{\omega - \Omega(\underline{x},t)}{\frac{\omega}{2} + (\omega - \Omega(\underline{x},t))}, & \Omega(\underline{x}, t) > \omega_{th} \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

where

$$\Omega(\underline{x}, t) = \frac{\partial \theta}{\partial t},$$

and ω_{th} represents a frequency threshold below which the amplitude is set to zero. The synthetic signal is then given by

$$s(\underline{x}, t) = \frac{1}{2} r(\underline{x}, t) \left(1 + \cos(\theta(\underline{x}, t))\right) + \xi_\sigma(t). \tag{10}$$

2.1.2 Numerical implementation

Equations (4)–(8) were solved numerically in a Cartesian coordinate system. The spatial domain was discretised using a regular square lattice with spatial step $\Delta x = \Delta y$. Spatial operators were approximated using a central difference scheme given by

$$\begin{aligned} \frac{\partial u}{\partial x} &\approx \frac{1}{2\Delta x}(u(x_{i+1}, y_j) - u(x_{i-1}, y_j)), \\ \frac{\partial u}{\partial y} &\approx \frac{1}{2\Delta x}(u(x_i, y_{j+1}) - u(x_i, y_{j-1})), \\ \nabla^2 u &\approx \frac{1}{\Delta x^2}(u(x_{i+1}, y_j) + u(x_{i-1}, y_j) \\ &\quad + u(x_i, y_{j+1}) + u(x_i, y_{j-1}) \\ &\quad - 4u(x_i, y_j)). \end{aligned} \tag{11}$$

The time derivative we approximated using a Forward Euler scheme.

2.2 Phase reconstruction

Let $s(\mathbf{x}, t)$ represent a spatio-temporal oscillatory signal. The steps used to generate the reconstructed phase profile are outlined below.

Parameter	Value	Definition	Unit
A	4	Coupling constant	c.d$^2h^{-1}$
B	1	Coupling constant	c.d$^2h^{-1}$
ω	2.5	Oscillation frequency	h^{-1}
ω_{th}	0.25	Oscillation threshold	h^{-1}
t_{seg}	16	Segmentation time	h
k_1	3	Amplitude modulation	h
R	20	Domain radius	c.d.
Δx	1	Spatial step	c.d.
Δt	0.025	Time step	h

Table 1: A table with parameters values used in synthetic data. c.d. cell diameter (10μm).

2.2.1 Multivariate Empirical Mode Decomposition (MEMD)

A multivariate signal $\mathbf{s}(t_k) = [x_1(t_k), x_2(t_k), ...]$ is decomposed by the Multivariate Empirical Mode Decomposition as follows:

1. The signal $\hat{\mathbf{s}}(t_k)$ is defined to be the original signal $\mathbf{s}(t_k)$.
2. A suitable pointset (n_d directions) is chosen on on a $(n-1)$-sphere.
3. The multivariate signal $\mathbf{s}(t_k)$ is projected along the direction vector $\mathbf{x}(t_k)$ to obtain the projection $p^{\varphi_\gamma}(t)$.
4. The time instants $t_i^{\varphi_\gamma}$ for each maxima of the set of projected signals are detected.
5. The multivariate envelope $e^{\varphi_\gamma}(t_k)$ is computed by interpolating $[t^{\varphi_\gamma}, \mathbf{s}(t^{\varphi_\gamma})]$.
6. For a set of W direction vectors, the mean is computed as
$$\mathbf{m}(t_k) = \frac{1}{W}\sum_{\gamma}^{W} e^{\varphi_\gamma}(t_k).$$
7. The mean is subtracted to obtain the detail
$$\mathbf{d}(t_k) = \mathbf{s}(t_k) - \mathbf{m}(t_k).$$
8. The variance of the detail is computed
$$S = \sum_{t_k}\sum_{N}(\mathbf{m}(t_k))^2.$$
9. If S is smaller than the stopping criterion, the n-th IMF is defined to be $\mathbf{d}(t_k)$ iand the process is repeated with $\hat{\mathbf{s}}(t_k)$ to be $\hat{\mathbf{s}}(t_k) - \mathbf{d}(t_k)$. Otherwise, $\hat{\mathbf{s}}(t_k)$ is define to be $\mathbf{d}(t_k)$ and Steps 3-8 are repeated.

10. The process is iterated until the remaining signal $\hat{\mathbf{s}}(t_k)$ does not have detectable extrema, which is defined to be $\mathbf{D}(t_k)$.

2.2.2 Spatially localised Multivariate Empirical Mode Decomposition (SLMEMD)

Let $s_{ij}(t)$ represent the signal at the ij^{th} point on a regular lattice. A multivariate signal $\bar{s}(\underline{\chi}, t)$ is defined with components that comprise the signal at the ij^{th} lattice point and nearest neighbours, i.e.

$$\bar{s}_{ij}(t) = [s_{ij}(t)s_{i+1j}(t)s_{i-1j}(t)s_{ij+1}(t)s_{ij-1}(t)]. \tag{12}$$

Multivariate Empirical Mode Decomposition (MEMD) is applied to the multivariate signal $\bar{s}_{ij}(t)$, as described above, yielding the decomposition

$$\bar{s}_{ij}(t) = \sum_{k=1}^{N} IMF_{ij,k}(t)$$

High frequency noise is eliminated by eliminating high frequency modes, i.e.

$$\hat{s}_{ij}(t) = \sum_{k=k_1}^{N} IMF_{ij,k}(t) \tag{13}$$

Due to (slow oscillating) trends that hinder the application of the Hilbert Transform (Pikovsky et al., 2001), the Empirical Mode Decomposition (Huang et al., 1998) is applied to the residual, i.e.

$$\hat{s}_{ij}(t) = \sum_{k=k_1}^{N} \overline{IMF}_{ij,k}(t)$$

and the reconstructed signal is defined to be the first IMF, i.e.

$$\tilde{s}_{ij}(t) = \overline{IMF}_{ij,1}(t) \tag{14}$$

Application of the Hilbert Transform, $\mathcal{H}$, yields the analytic signal

$$\tilde{s}_{ij}(t) + i\mathcal{H}\left\{\tilde{s}_{ij}(t)\right\} = r(t)e^{i\theta(t)}, \tag{15}$$

where $r(t)$ and θ are the instantaneous reconstructed amplitude and phase respectively. The reconstructed instantaneous frequency is computed as

$$\omega = \frac{d\theta}{dt}. \tag{16}$$

The above process is applied to all voxels in the sample.

2.2.3 Identifying the oscillatory domain

To track the oscillatory region of the mPSM explant, a time-dependent mask is constructed. At a given time point, a mask is defined by thresholding the signal, i.e.

$$m(\mathbf{x}, t; \gamma) = s(\mathbf{x}, t) > \gamma, \tag{17}$$

where γ is a threshold parameter that is optimised on a sample-by-sample basis.

To compute the boundary of the oscillatory domain, erosion is applied on the mask, i.e.

$$m(\mathbf{x}, t; \gamma)' = m(\mathbf{x}, t; \gamma) \ominus K_d, \tag{18}$$

where K_d is a kernel of radius R_d and $\ominus$ is the erosion operator, and the boundary, which is the eroded region, is given by

$$w(\mathbf{x}, t; \gamma) = m(\mathbf{x}, t, \gamma) - m(\mathbf{x}, t, \gamma)'. \tag{19}$$

To identify the optimal value of the threshold, W_γ is defined to be the sum of the intensity of all voxels on the boundary, i.e.

$$W_\gamma = \sum_i \sum_j w(\mathbf{x}, t; \gamma) \otimes s(\mathbf{x}, t) \tag{20}$$

As in the experimental data oscillations are arrested after they achieve maximum amplitude, the optimal value of γ, γ_{opt}, is defined to be

$$\gamma_{opt} = \max_\gamma [W_\gamma]. \tag{21}$$

To eliminate potential holes, both dilation and erosion are applied, i.e.

$$M(\mathbf{x}, t; \gamma_{opt}) = (m(\mathbf{x}, t; \gamma_{opt}) \oplus K_{ed}) \ominus K_{ed}, \tag{22}$$

where K_{ed} is a kernel of radius R_{ed} and $\oplus$ is the binary dilation operator. Finally the mask is time-averaged over three time points. Applying the mask $M(\mathbf{x}, t; \gamma_{opt})$ to the phase yields the oscillatory region of the PSM explant, i.e.

$$\theta(\mathbf{x}, t) = \theta(\mathbf{x}, t) \otimes M(\mathbf{x}, t; \gamma_{opt}).$$

2.2.4 Phase on the boundary

During the segmenting phase of mPSM explants, the oscillating domain reduces in size. Additionally, the signal intensity is largest on the boundary of the tissue and there is no signal outside of the actively oscillating domain. However, the pattern has information (most peripheral region have experienced fewer oscillations than those in the interior).

To retain the phase information we identify regions of the mPSM explant where the signal is being lost, i.e. we identify points in space and time, $(\underline{x}^*, t^*)$, that satisfy

$$M(\mathbf{x}, t + \Delta t; \gamma_{opt}) - M(\mathbf{x}, t; \gamma_{opt}) = 1 \tag{23}$$

and impose

$$\theta(\underline{x}^*, t) = \theta(\underline{x}^*, t^*), \qquad for \quad t \geq t^*. \tag{24}$$

2.2.5 Initialisation and unwrapping of phase

The unwrapped phase is defined to be

$$\theta_U(\underline{x},t) = \theta(\underline{x},t) + 2n(\underline{x},t)\pi,$$

where $n(\underline{x},t)$ represents the number of complete cycles that a voxel has gone through in a particular point in space and time.

In general it is not possible to infer $n(\mathbf{x},0)$ using the raw signal (one does not know how many cycles have elapsed at the beginning of the recording).

To define an unwrapped phase in the case of mPSM explants, we use the fact that there is a maximum phase drop of 2π across the PSM tissue. Hence a time t^* is identified where a region of mPSM tissue is just entering a new cycle. Hence it is approximated that

$$n(\underline{x},t^*) = 0.$$

To compute the unwrapped phase profile $n(\mathbf{x},\mathbf{t})$ for $t > 0$, $\pm 2\pi$ jumps are identified using the function `unwrap` implemented in `MATLAB`).

2.2.6 Smoothing of the phase profile

We compute a median filter of radius R_{filt} voxels.

2.2.7 Computing sample trajectories

Let $\mathbf{x}_0$ be an initially selected point in the oscillatory domain. The phase gradient, $\nabla\theta$, is computed and a gradient descent algorithm is used to iteratively compute the trajectory

$$\mathbf{x}_i = \mathbf{x}_{i-1} - k_{gd}\frac{\nabla\theta}{|\nabla\theta|}, \quad i = 1,..,N_{gd}.$$

where the arc length is given by

$$\sigma_i = \sum_{j=1}^{i} |\mathbf{x}_j - \mathbf{x}_{j-1}|,$$

k_{gd} is the step size and N_{gd} is the maximum number of steps.

The phase on the trajectory is given by

$$\theta(\sigma_i,t) \quad i = 1,..,N.$$

We apply a smoothing cubic spline that identifies the smoothed phase, $\theta_S(\sigma_i,t)$, to be the function that minimises

$$\alpha\sum_{i=1}^{n} |\ \theta_S(\sigma_i,t) - \theta(\sigma_i t)\ |^2 + (1-\alpha)\int \left|\nabla^2\overline{\theta}_S(\sigma,t)\right|^2 d\sigma, \tag{25}$$

where $\alpha \in [0, 1]$ a regularisation parameter. The phase gradient is approximated as

$$\frac{\theta_U(\sigma_i) - \theta_U(\sigma_{i-1})}{\sigma_i - \sigma_{i-1}}.$$

2.3 Metrics

2.3.1 Wave tracking

To measure tissue scale periodicity, the fraction of cells that have undergone k cycles at a given time t is computed. We define the number of cycles in the actively oscillating domain to be

$$\eta(\underline{x}, t) = n(\underline{x}, t) \otimes M(\underline{x}, t).$$

Defining the number of actively oscillating voxels at time t to be

$$V(t) = \sum_i M(\underline{x}_i, t),$$

the fraction of voxels with k elapsed cycles is given by

$$w_k(t) = \frac{1}{V(t)} \sum_i h(\eta(x_i, t) - k), \tag{26}$$

and $h(.)$ is an indicator function given by

$$h(x) = \begin{cases} 1, & x = 0, \\ 0, & x \neq 0. \end{cases} \tag{27}$$

The time of the emergence of the k^{th} wave, t_k^*, is defined to be

$$w_k(t_k^*) = \chi_0, \tag{28}$$

where χ_0 is a constant. The tissue scale period is defined to be the time between the emergence of two consecutive waves, i.e.

$$T_k = t_k^* - t_{k-1}^*. \tag{29}$$

2.3.2 Instantaneous frequency

Suppose that $SG_n(x_i, y_j, t_k, \omega)$ is a time frequency spectrogram measured at voxel x_i,y_j is the n^{th} data sample. The spectrogram of a sample is given by

$$SG_{\tau,n}(t, \omega) = \sum_i \sum_j SG_n(x_i, y_j, t_k, \omega). \tag{30}$$

Parameter	Value	Definition
χ_0	0.01	Mask threshold
α	0.05	Spline smoothing
R_{filt}	3	Median filter radius
k_{gd}	1.2	Gradient descent step size
N_{gd}	20	Max steps in gradient descent
ω_{ms}	0.53	Mode selection frequency threshold
R_{ed}	3	Mask erosion kernel radius
R_d	5-10	Mask erosion kernel radius
n_D	300	Number of direction vectors in MEMD

Table 2: A table with parameters values used in phase reconstruction and metric evaluation.

2.3.3 Differentiation rate

Defining $V(t)$ to be the number of actively oscillating voxels at time t, the averaged normalised differentiation rate is defined to be

$$v = \frac{1}{V(t)} \frac{dV}{dt}. \tag{31}$$

2.4 Experimental

2.4.1 Mouse line

The *LuVeLu* mouse (*Mus musculus*) (a gift from O. Pourquie) expresses *Venus-YFP* under the control of a 2.1kb fragment of the *Lunatic Fringe* (*Lfng*) promoter. The mRNA contains the 3'UTR of *Lfng* and the protein is fused to a PEST domain, to destabilise both the RNA and the protein and ensure clear oscillations (see Aulehla et al. (2008) for further details

2.4.2 Mating procedure

Mouse E10.5 embryos were generated and the line was maintained by crossing *LuVeLu* males to stock CD1 females, as the *LuVeLu* construct is lethal in homozygotes.

2.4.3 Mouse genotyping

To renew the mouse line, timed matings were performed and the litters genotyped to ensure the *LuVeLu* construct is still present. A diagnostic PCR was performed on total DNA from an ear biopsy tissue of individual animals. DNA was extracted by incubation in microLYSIS-Plus buffer (*Thistle Scientific*) through the following PCR protocol: 65°C for 15 minutes, 96°C for 2 minutes, 65°C for 4 minutes, 96°C for 1 minute, 65°C for 1 minute, 96°C for 30 seconds and 8°C until stopped.

The PCR mix was generated by adding 2μl of the lysed solution to a solution containing 1.25μl GoTaq™Flexy polymerase (*Promega*), 0.31 mM dNTPs (*Promega*), 1.25 mM $MgCl_2$, 1 GoTaq™Flexi PCR buffer (*Promega*) and 20 pmol of each of the following four primers: (i) Ala1 (Forward), 5'-tgctgctgcccgacaaccact-3'; (ii) Ala3 (Reverse), 5-tgaagaacacgactgcccagc-3; (iii) IMR0015, 5'-caaatgttgcttgtctggtg-3; (iv) IMR0016, 5-gtcagtcgagtgcacagttt-3. Distilled water was added to the solution to reach a final volume of 20μl. The following PCR protocol was used: 94°C for 2 minutes; 35 cycles of [92°C for 45 seconds, 59°C for 40 seconds, 72°C for 40 seconds]; 75°C for 5 minutes; and 4°C until stopped.
PCR samples were analysed through electrophoresis, by loading 5μl of each PCR product onto a 1% agarose gel with 1:10,000 Gel Red™(*Biotium/VWR*®) and run for 20 minutes at 100 volts. The result was visualised using an UV light box. Wild type CD1 presented a single fragment of 200 bp whilst $LuVeLu^{+}/-$ presented this fragment together with a 461 bp fragment.

2.4.4 *ex vivo* culture system

A 35 mm FluoroDish (*World Precission Instruments*™) with coverglass bottom was coated with a 50 μg/ml fibronectin (*Sigma*) in a 100 mM sodium chloride (NaCl) solution made with double distilled water (ddH_2O) before the dissection. The dish was incubated in the solution either for 4 hours at room temperature or overnight at 4 °C with agitation. The solution was later removed and the dish left until it was completely dried, around 30 minutes.
The dish was washed in tail bud culture medium of DMEM/F12 with no phenol red (*Gibco/Life Technologies*™) with 0.5 mM glucose (*Sigma*), 2mM L-glutamine (*Gibco/Life Technologies*™), 1% bovine serum albumin (BSA) (*Sigma*), penicillin/ streptomycin(*Gibco/Life Technologies*™) for 30 minutes at room temperature.
Embryos were harvested at E10.5 from timed matings. Individual embryos were taken from the uterine horn in sterile PBS using forceps and transferred to dissection media. To identify *luVeLu* positive embroys, tails were cut and transferred to a pre-warmed tail bud dissection media (tail bud culture media + 10 mM HEPES (*Sigma*) in a multi-well dish.
After *LuVeLu* tail identification, the tail bud was isolated from each tail posterior to the neuropore and transferred to an imaging disk, with the cut facing downwards, towards the fibronectin-coated surface.
Explants were incubated at 38.5°C, 5% CO_2 for 1 hour to allow the explant to adhere to the fibronectin-coated surface, before live imaging. Embryos were transferred to a confocal microscope, as described in Section 2.4.5, and imaged for 24 hours at 37°C, 5% CO_2 and ambient O_2 levels .

2.4.5 Imaging

Tail bud explants were prepared as described above. The imaging dish was transferred to a 37°C heated stage with a heated chamber at 37°C with 5% CO_2 and ambient O_2 of a Zeiss 710 inverted confocal microscope. Explants

were imaged using a EC Plan-Neoufluar 10x/0.30 dry objective (*Zeiss*, M27). The *LuVeLu* fluorophore was excited by using a 514nm Argon laser.
Samples were scanned bi-directionally, to increase acquisition speed, and averaged 8 times per line, with a spatial resolution of 1024x1024 pixels and a temporal resolution of 15 minutes. Three optical planes were acquired at 14μm intervals and starting at the plane of the cut, and moving upwards.

3 Results

To obtain a synthetic dataset with phase dynamics that are qualitatively similar to mPSM explants, the numerical solution of equations (4)–(8) was computed and defined on a dynamic spatial domain given by equation (9) (see Figure 1 (a)). Figure 1 (b) depicts the dynamic domain (equation (9)), Figure 1 (c) depicts the phase (see Section 2.2.4) and in Figure 1 (d) depicts the unwrapped phase (solution of equations (4) – (8)). The goal in this section is to develop a methodology for inference of the unwrapped phase profile presented in Figure 1 (d) from the experimental-like signal presented in Figure 1 (a).

To reconstruct the phase profile for an oscillating signal defined on a dynamic domain, a phase reconstruction algorithm (see Section 2.2) was developed. The first step in the algorithm uses a novel extension to EMD that we call spatially localised empirical mode decomposition (SL-MEMD). SL-MEMD is a multivariate EMD in which a multivariate signal is defined that considers the signal in each voxel and its locally synchronised neighbours (see Section 2.2.2). After recovering the signal using SLMEMD (see Figure 2 (a)), a mask is defined that identifies the actively oscillating region of the signal (see Figure 2 (b)). To record the phase history in regions of the sample that have stopped oscillating, the phase is fixed in time as the boundary propagates inwards (see Figure 2 (c)). Finally, the reconstructed phase is unwrapped and smoothed using a median filter (see Figure 2 (d)).

To infer the tissue scale period, a wave tracking algorithm was developed that identifies the emergence of successive oscillatory waves (see Figure 3 (a) and Section 2.3.1). By computing the time that elapses between the emergence of successive waves (see Figure 3 (b)), the tissue scale oscillation frequency is accurately computed (see Figure 3 (c)). To describe the differentiation rate of the tissue we computed the normalised rate of change of area of the actively oscillating region (see Figure 3 (d)). To describe the frequency dynamics throughout the population, we computed an instantaneous frequency spectrogram (see Figure 3 (d) and Section 2.3.2).

The generation of kymograph requires the use to specify a spatial axis upon which dynamics are quantified. We have developed a gradient descent method which define trajectories that are locally tangential to the phase gradient (see Figure 3 (f)). In Figure 3 (e) we plot the phase dynamics as a function of arc length along five such trajectories. In Figure Figure 3 (f) we plot the oscillation frequency against phase gradient. Note that the underlying structure of the

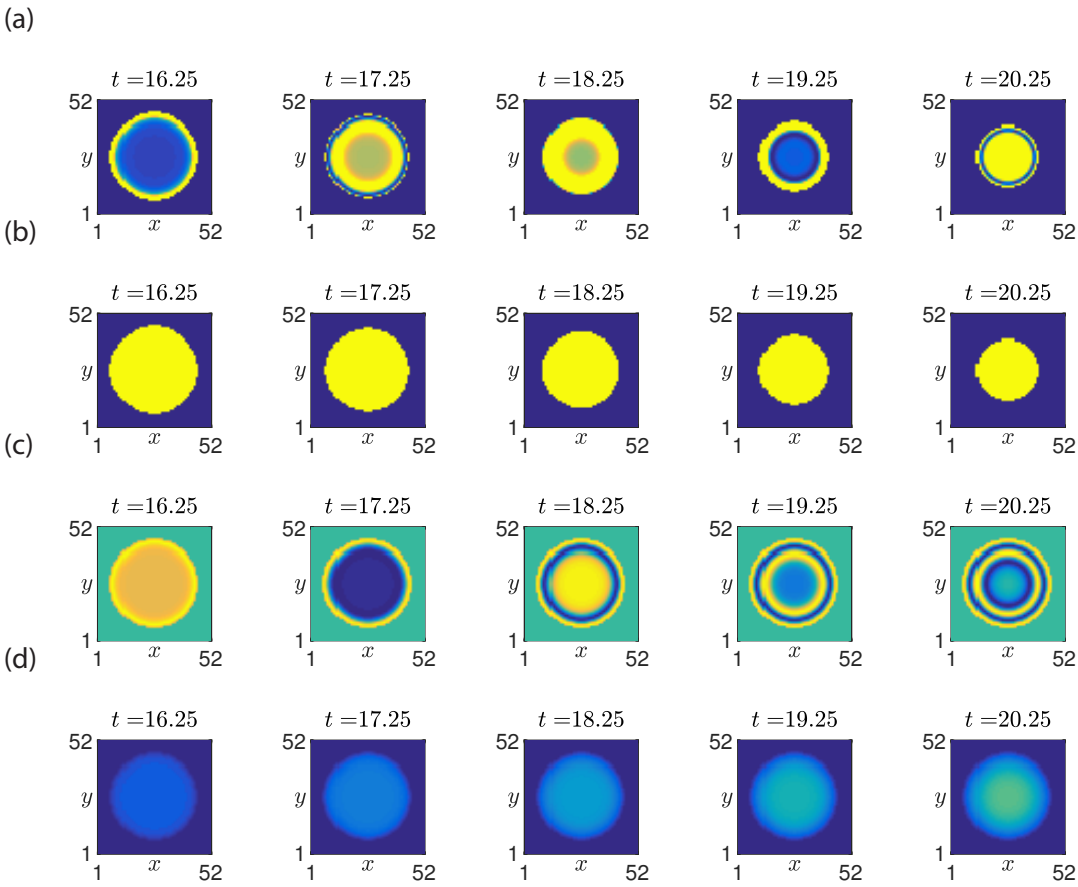

Figure 1: The synthetic dataset $s(x, y, t)$. (a) Snapshots of the synthetic signal (equation (10)). (b) Snapshots of the actively signalling domain (equation (9)). (c) Snapshots of the sine of the reconstructed oscillator phase, $\hat{\theta}$. (d) Snapshots of the oscillator phase, $\hat{\theta}$. See Table 2 for parameter values.

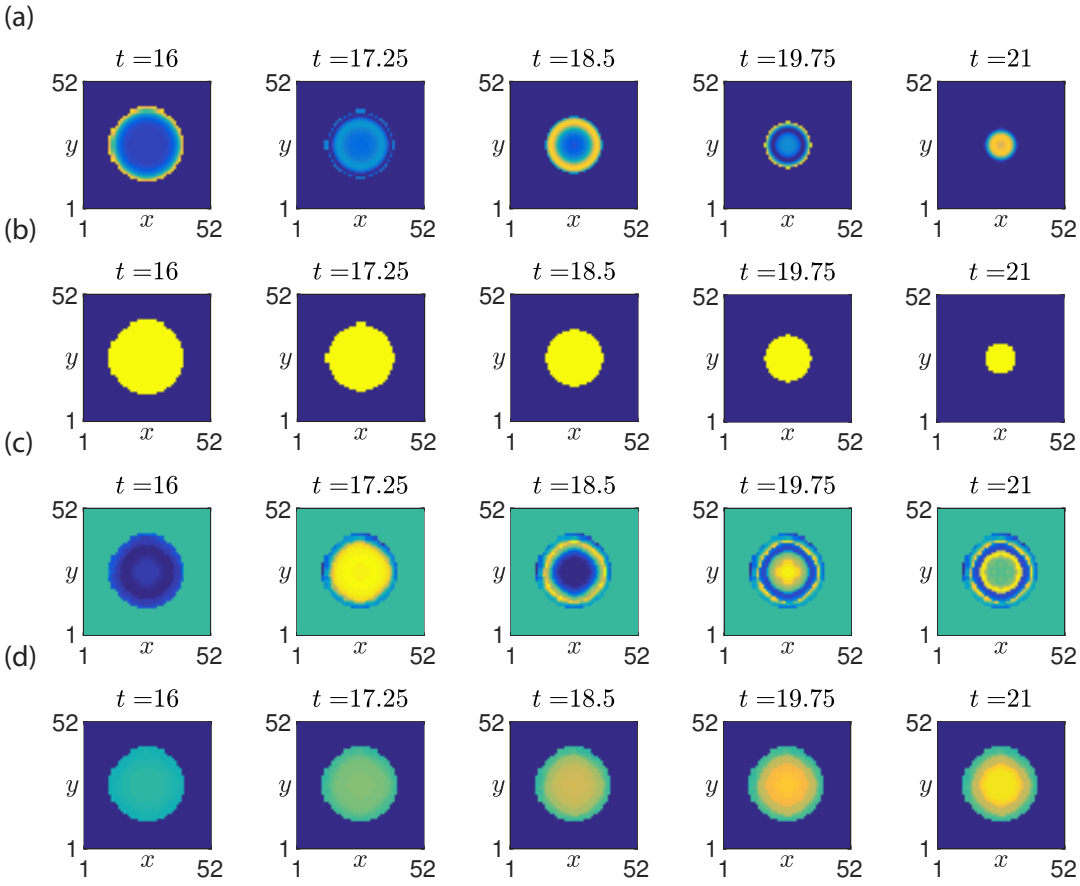

Figure 2: Phase reconstruction of the synthetic dataset $s(x, y, t)$. (a) Snapshots of the synthetic signal. (b) Snapshots of the mask that defined the extent of the actively signalling domain. (c) Snapshots of the sine of the reconstructed oscillator phase, $\hat{\theta}$. (d) Snapshots of the oscillator phase, $\hat{\theta}$. See Table 2 for parameter values.

model used to generate the synthetic data is recovered (i.e. the oscillation frequency is a linear function of the other variables). These results show that the proposed methodology is capable of describing a range of features of phase dynamics in simulated mPSM explants.

To test the proposed phase reconstruction technique on experimental data, mPSM explants from a real time reporter of the mouse somitogenesis clock were cultured on fibronectin coated glass coverslips. We found that explants were viable and exhibited spatiotemporal oscillations of gene expression (see Figure 4 (a) and (b)). Moreover we observed a front of high reporter activity that propagates from the periphery towards the centre of the tissue and oscillatory waves of gene expression that propagate from the centre to the periphery. However, in our hands there was significant inter-sample variability and the geometry of the sample appeared to to play a dominant role in the pattern of gene expression.

To reconstruct the phase profile we applied the algorithm outline in Section 2.2. The actively oscillating region was determined (see Figure 4 (c)) and application of SLMEMD yielded the phase dynamics in the actively oscillating region of the signal. The full phase history of the sample was defined by recording the phase at the segmentation boundary (see Figure 4 (d) and (e)).

The wave tracking algorithm (see Section 2.3.1) was applied to the reconstructed phase profile and the tissue scale frequency of the sample was computed (see Figure 4 (a) - (c)). The average oscillation frequency was in close agreement with that measured in a previous study (Lauschke et al., 2013). To identify the time at which the sample changed from growing to segmenting phases, we computed the tissue differentiation rate (see Figure Figure 4 (d)) and showed that it was initially positive (the sample is expanding) but becomes negative at around 15 h (Lauschke et al., 2013).

To explore spatial structure of the reconstructed phase profile we used the phase gradient descent method to identify trajectories along which phase waves propagate (see Figure 5 (f)). By examining phase dynamics along these trajectories we compute automated kymographs that allow spatio-temporal phase dynamics to be systematically explored (see Figure 5 (g)). We find that in general the boundaries of the mPSM sample are not homogeneous (at a given point in time some regions of the boundary stop oscillating whilst others do not). We also identified that the the oscillation frequency varies inversely with the magnitude of the phase gradient (see Figure 5 (h)). Finally, along a given trajectory a scaling phenomenon whereby the phase gradient steepens as the size of the oscillatory domain decreases (see Figure 5 (h)).

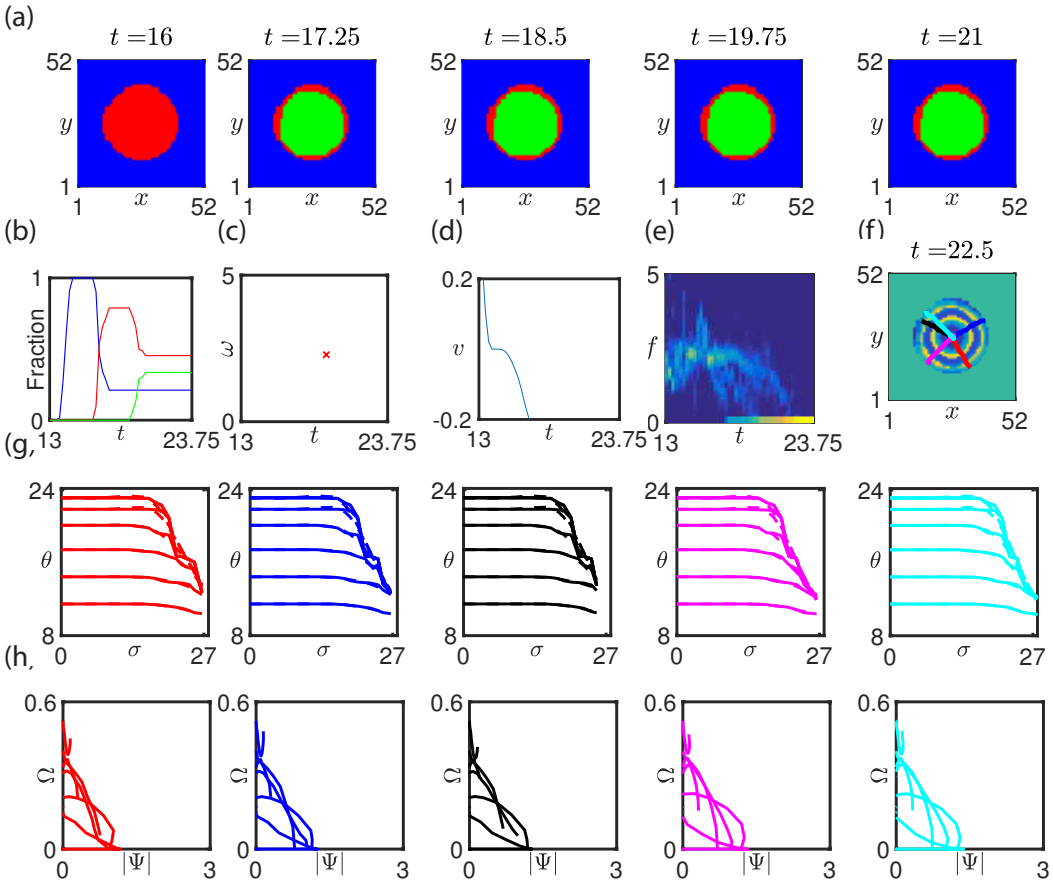

Figure 3: Application of metrics to reconstructed phase dynamics. (a) Snapshots of the emergence of successive waves. (b) The area spanned by each wave is plotted against time. (c) The inter-wave frequency is plotted against time (see equation (29)). (d) The tissue differentiation rate is plotted against time (equation (31)). (e) The distribution of instantaneous frequencies is plotted against time. (f) Identification of trajectories using gradient descent. (g) Phase is plotted against arc length, σ, along different gradient descent trajectories. (h) The oscillation frequency is plotted against the phase gradient along different gradient descent trajectories.

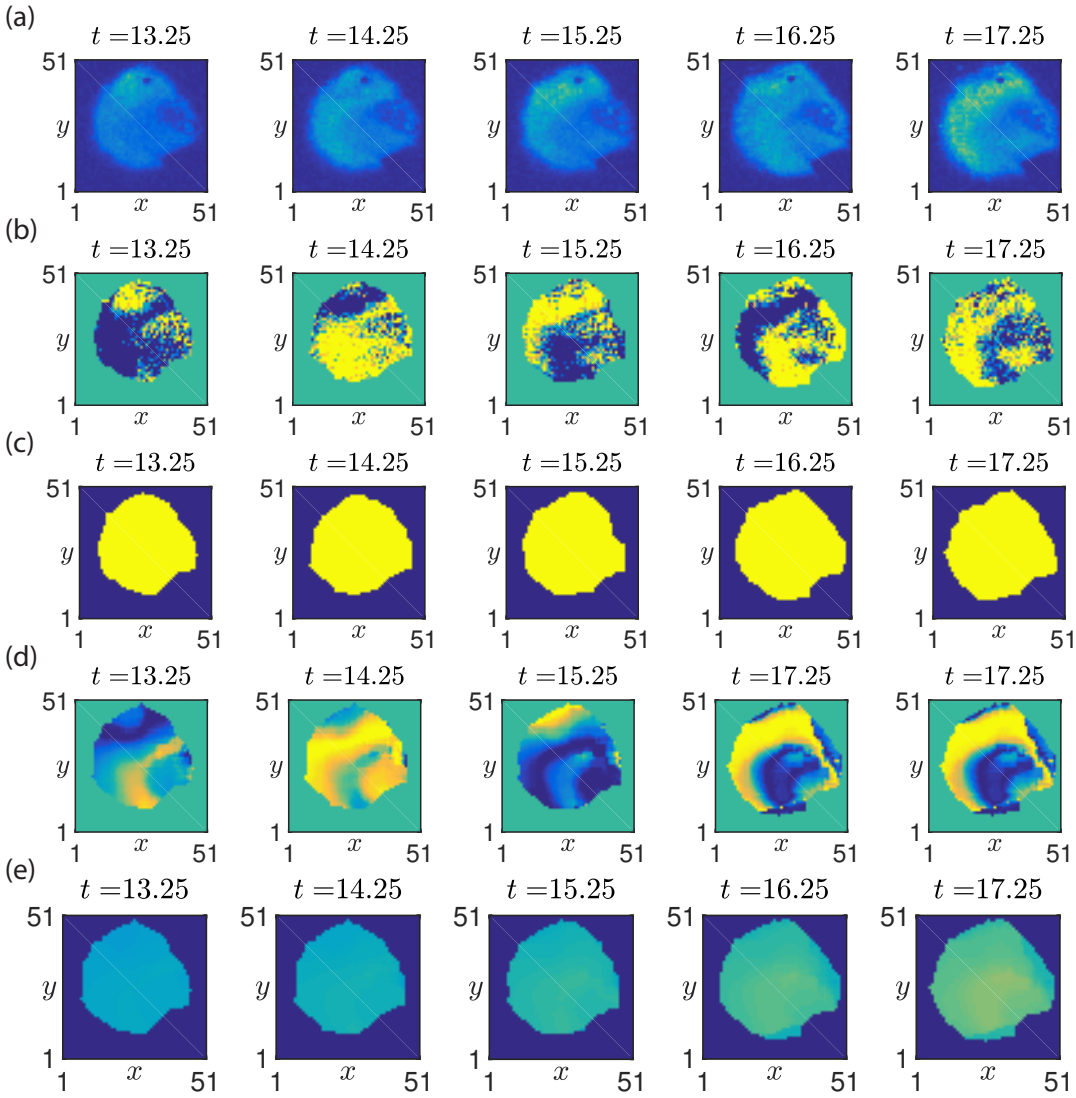

Figure 4: Recovered spatiotemporal dynamics from an mPSM explant. (a) Snapshots of the fluorescent signal from the LuVeLu reporter. (b) Snapshots of the fluorescent signal after application of a moving average filter. (c) Snapshots of the mask that defined the extent of the actively signalling domain. (d) Snapshots of the sine of the reconstructed oscillator phase, $\hat{\theta}$. (e) Snapshots of the oscillator phase, $\hat{\theta}$. See Table 2 for parameter values.

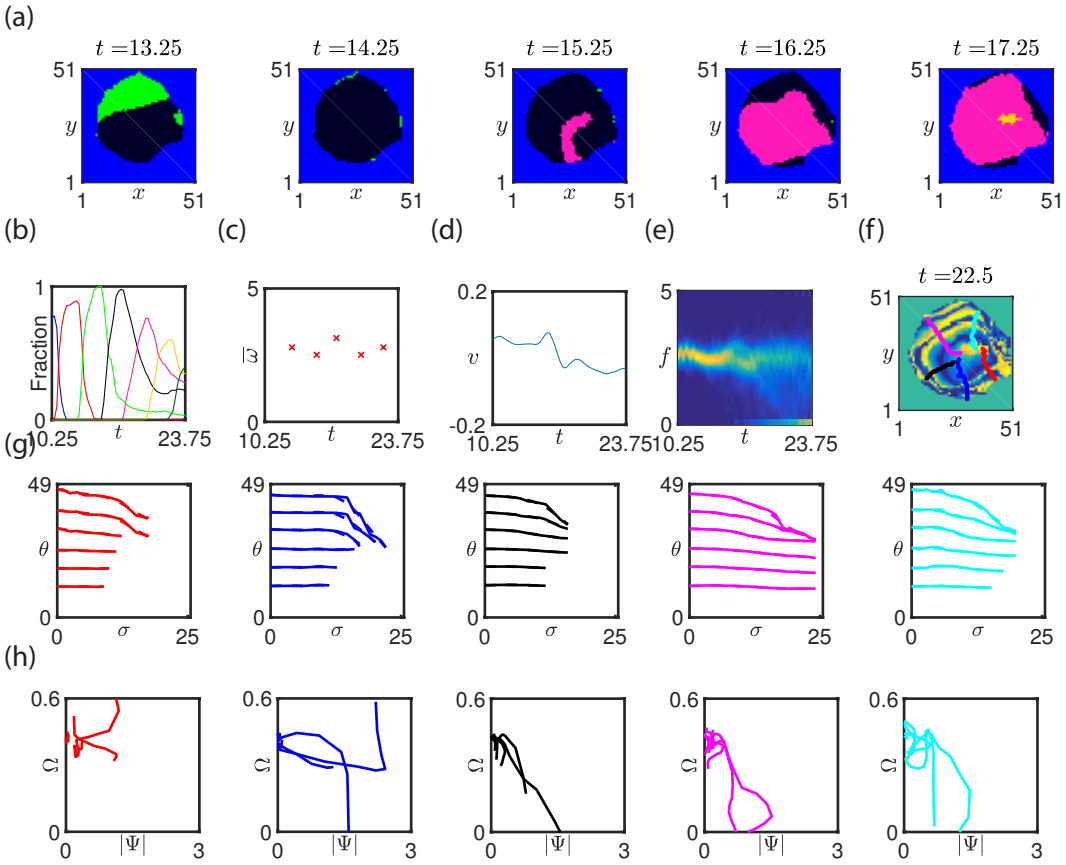

Figure 5: Quantitative analysis of mPSM explants. (a) Snapshots of the emergence of successive waves. (b) The area spanned by each wave is plotted against time. (c) The inter-wave frequency is plotted against time (see equation (29)). (d) The tissue differentiation rate is plotted against time (equation (31)). (e) The distribution of instantaneous frequencies is plotted against time. (f) Identification of trajectories using gradient descent. (g) Phase is plotted against arc length, σ, along different gradient descent trajectories. (h) The oscillation frequency is plotted against the phase gradient along different gradient descent trajectories.

4 Discussion

Spatio-temporal oscillations play a crucial role in many biological systems (e.g. neural signalling, cardiac waves, calcium waves, somitogenesis). The quantification of the features of oscillations provides an information-rich means of characterising system behaviour as well as a bridge between experimental observation and theory.

The phase of an oscillator, a variable that describes relative progression through a cycle, allows the phenomenological features of its oscillatory behaviour to be described. There are numerous methods available to define phase (e.g. Fourier, Wavelet, Hilbert, EMD), each of which can yield robust phase reconstructions in certain contexts. Data-driven methods, such as EMD, are well-suited to the study of nonstationary, nonlinear signals.

During development of the vertebrate embryo, the presomitic mesoderm sequentially segments into pairs of somites at regular intervals in time. Underlying temporal periodicity of somite formation is a molecular oscillator, known as the segmentation clock, that exhibits striking spatio-temporal patterns of gene expression. Recent advances in the development of real-time reporters of gene expression (e.g. Soroldoni and Oates, 2011; Aulehla et al., 2008; Masamizu et al., 2006) have resulted in the identification of novel phenomena, such as aDoppler effect and phase gradient scaling, as well the generation of large datasets that contain complex spatio-temporal patterns. There is not yet a consensus on which methods are optimal for quantifying phenomenology of the spatio-temporal dynamics.

In this study we have developed a phase reconstruction methodology based on empirical model decomposition that allows inference of phase dynamics on a dynamic spatial domain. The methodology uses a variant of EMD, which we call SLMEMD, to infer phase dynamics. Moreover, we develop a set of metrics that when applied to the reconstructed phase profile yield the quantification of biologically meaningful variables.

In order to validate the phase reconstruction methodology, we defined a synthetic dataset that mimicked many known features of segmentation clock dynamics. We showed that the ground truth phase could be accurately reconstructed using the proposed methodology and that a range of biologically relevant quantities (e.g. tissue scale period, instantaneous frequency distribution, geometrical features) can be accurately inferred from the reconstructed phase profile.

To determine if the phase construction methodology could be applied to experimental measurements form a real time reporter of the segmentation clock, we performed a series of experiments in which mPSM explants from the LuVeLu reporter mouse were cultured as monolayers on fibronectin-coated glass slides. We inferred the phase dynamics in the mPSM sample using the phase reconstruction methodology and, consistent with previous observations, found: (i) two distinct phases of behaviour (expanding and segmenting); (ii) outwardly propagating oscillatory waves of gene expression; and (iii) an inwardly propagating wavefront that defines a peripheral limit of the oscillatory domain. Ap-

plication of the metrics yielded quantitative results in good agreement with the previous observations. However, by recovering phase dynamics in the full spatial domain, geometrical features of the phase profiles are described.

We have proposed SLMEMD, a novel variant on EMD that is applicable to locally synchronised spatio-temporally oscillating systems. The defining property of SL-MEMD, which is that the signal in neighbouring voxels is used in the inference of phase at a given voxel, increases the robustness of phase inference compared with EMD but does not require the tuning of noise strength as is the case with NAMEMD. We note that we have reproduced our result using NAMEMD and have not found significant deviation between the methods for the signals that we have considered.

EMD and its derivative methods are empirical. To deal with this property, we have developed synthetic datasets that approximate the experimental data from which we would like to infer phase dynamics. Whilst we have explored perturbations round these datasets, we note that aspects of the methodology may need to be optimised for sufficiently different problems.

One of the advantages of the phase reconstruction approach adopted in this study is that the phase dynamics are reconstructed in the full spatial domain. Compared with kymograph analyses, which are frequently used in the literature and require the arbitrary specification of a spatial axis upon which to analyse the signal, the proposed approach allows one to quantify the evolving geometry of phase profile.

A limitation of the current study is that individual cells are not tracked. In principle, artefacts in the phase dynamics could be induced by cell motion. In the mPSM explant analysed in this study, we considered only the segmenting phase ($t > 16h$) during which cell motion is greatly reduced.

An important caveat with this work is that the experimentally measured signal is periodic. Hence at given instant in time one sees a readout of gene expression but not the absolute number of cycles that have elapsed. This issue is problematic in the definition of the unwrapped phase as it is impossible to infer the absolute phase the reporter signal. However, mPSM explants exhibit approximately spatially homogeneous patterns of gene expression at early times, a property that is used to define the initial condition for the unwrapped phase.

In this study we have developed a methodology for reconstructing phase dynamics in mPSM explants. In future work we will use the phase reconstruction techniques to investigate the response of mPSM explants to chemical perturbation and we will explore the extent to which low dimensional mathematical models can be used to reproduce observed phenomenology.

References

Javier Alegre-Cortés, Cristina Soto-Sánchez, ÁG Pizá, Ana Lia Albarracín, Fernando Daniel Farfán, Carmelo Jose Felice, and Eduardo Fernández. Time–frequency analysis of neuronal populations with instantaneous resolution

based on noise-assisted multivariate empirical mode decomposition. *Journal of neuroscience methods*, 267:35–44, 2016.

Alexander Aulehla, Winfried Wiegraebe, Valerie Baubet, Matthias B Wahl, Chuxia Deng, Makoto Taketo, Mark Lewandoski, and Olivier Pourquié. A β-catenin gradient links the clock and wavefront systems in mouse embryo segmentation. *Nature Cell Biology*, 10(2):186, 2008.

Sharif MA Bhuiyan, Reza R Adhami, and Jesmin F Khan. Fast and adaptive bidimensional empirical mode decomposition using order-statistics filter based envelope estimation. *EURASIP Journal on Advances in Signal Processing*, 2008(1):728356, 2008.

Chih-Sung Chen and Yih Jeng. Two-dimensional nonlinear geophysical data filtering using the multidimensional eemd method. *Journal of Applied Geophysics*, 111:256–270, 2014.

S H R Davies and C J James. Using empirical mode decomposition with spatio-temporal dynamics to classify single-trial motor imagery in bci. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4631–4634, 2014.

Emilie A Delaune, Paul François, Nathan P Shih, and Sharon L Amacher. Single-cell-resolution imaging of the impact of Notch signaling and mitosis on segmentation clock dynamics. *Developmental Cell*, 23(5):995–1005, 2012.

Richard Harang, Guillaume Bonnet, and Linda R Petzold. Wavos: a matlab toolkit for wavelet analysis and visualization of oscillatory systems. *BMC Research Notes*, 5(1):163, 2012.

Zhi He, Jun Li, Lin Liu, and Yi Shen. Three-dimensional empirical mode decomposition (temd): A fast approach motivated by separable filters. *Signal Processing*, 131:307–319, 2017.

Seth M Hirsh, Bingni W Brunton, and J Nathan Kutz. Data-driven spatiotemporal modal decomposition for time frequency analysis. *arXiv preprint arXiv:1806.08739*, 2018.

Norden E Huang, Zheng Shen, Steven R Long, Manli C Wu, Hsing H Shih, Quanan Zheng, Nai-Chyuan Yen, Chi Chao Tung, and Henry H Liu. The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 454, pages 903–995. The Royal Society, 1998.

Alexis Hubaud, Ido Regev, L Mahadevan, and Olivier Pourquie. Excitable dynamics and Yap-dependent mechanical cues drive the segmentation clock. *Cell*, 171(3):668–682, 2017.

Kazuyoshi Kikuchi and Bin Wang. Spatiotemporal wavelet transform and the multiscale behavior of the madden–julian oscillation. *Journal of Climate*, 23 (14):3814–3834, 2010.

Volker M Lauschke, Charisios D Tsiairis, Paul François, and Alexander Aulehla. Scaling of embryonic patterning based on phase-gradient encoding. *Nature*, 493(7430):101, 2013.

Stéphane Mallat. *A wavelet tour of signal processing*. Academic press, 1999.

Yoshito Masamizu, Toshiyuki Ohtsuka, Yoshiki Takashima, Hiroki Nagahara, Yoshiko Takenaka, Kenichi Yoshikawa, Hitoshi Okamura, and Ryoichiro Kageyama. Real-time imaging of the somite segmentation clock: revelation of unstable oscillators in the individual presomitic mesoderm cells. *Proceedings of the National Academy of Sciences*, 103(5):1313–1318, 2006.

Marina Matsumiya, Takehito Tomita, Kumiko Yoshioka-Kobayashi, Akihiro Isomura, and Ryoichiro Kageyama. ES cell-derived presomitic mesoderm-like tissues for analysis of synchronized oscillations in the segmentation clock. *Development*, 145(4), 2018.

Philip J Murray, Philip K Maini, and Ruth E Baker. The clock and wavefront model revisited. *Journal of Theoretical Biology*, 283(1):227–238, 2011.

Jean Claude Nunes, Yasmina Bouaoune, Eric Delechelle, Oumar Niang, and Ph Bunel. Image analysis by bidimensional empirical mode decomposition. *Image and Vision Computing*, 21(12):1019–1026, 2003.

Isabel Palmeirim, Domingos Henrique, David Ish-Horowicz, and Olivier Pourquié. Avian hairy gene expression identifies a molecular clock linked to vertebrate segmentation and somitogenesis. *Cell*, 91(5):639–648, 1997.

Cheolsoo Park, David Looney, Naveed ur Rehman, Alireza Ahrabian, and Danilo P Mandic. Classification of motor imagery bci using multivariate empirical mode decomposition. *IEEE Transactions on neural systems and rehabilitation engineering*, 21(1):10–22, 2013.

Arkady Pikovsky, Michael Rosenblum, Jurgen Kurths, and Jürgen Kurths. *Synchronization: a universal concept in nonlinear sciences*, volume 12. Cambridge university press, 2001.

Naveed Rehman and Danilo P Mandic. Multivariate empirical mode decomposition. In *Proceedings of The Royal Society of London A: Mathematical, Physical and Engineering Sciences*, page rspa20090502. The Royal Society, 2009.

Naveed Rehman and Danilo P Mandic. Empirical mode decomposition for trivariate signals. *IEEE Transactions on Signal Processing*, 58(3):1059–1068, 2010.

Naveed Rehman, Cheolsoo Park, Norden E Huang, and Danilo P Mandic. Emd via memd: multivariate noise-aided computation of standard emd. *Advances in Adaptive Data Analysis*, 5(02):1350007, 2013.

Jamal Riffi, Adnane Mohamed Mahraz, Abdelghafour Abbad, and Hamid Tairi. 3d extension of the fast and adaptive bidimensional empirical mode decomposition. *Multidimensional Systems and Signal Processing*, 26(3):823–834, 2015.

Gabriel Rilling, Patrick Flandrin, Paulo Gonçalves, and Jonathan M Lilly. Bivariate empirical mode decomposition. *IEEE signal processing letters*, 14(12): 936–939, 2007.

Jérémy Schmitt, Nelly Pustelnik, Pierre Borgnat, and Patrick Flandrin. 2d hilbert-huang transform. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 5377–5381. IEEE, 2014.

Katharina F Sonnen, Volker M Lauschke, Julia Uraji, Henning J Falk, Yvonne Petersen, Maja C Funk, Mathias Beaupeux, Paul François, Christoph A Merten, and Alexander Aulehla. Modulation of phase shift between Wnt and Notch signaling oscillations controls mesoderm segmentation. *Cell*, 172 (5):1079–1090, 2018.

Daniele Soroldoni and Andrew C Oates. Live transgenic reporters of the vertebrate embryo's segmentation clock. *Current Opinion in Genetics and Development*, 21(5):600–605, 2011.

Daniele Soroldoni, David J Jörg, Luis G Morelli, David L Richmond, Johannes Schindelin, Frank Jülicher, and Andrew C Oates. A Doppler effect in embryonic pattern formation. *Science*, 345(6193):222–225, 2014.

Charisios D Tsiairis and Alexander Aulehla. Self-organization of embryonic genetic oscillators into spatiotemporal wave patterns. *Cell*, 164(4):656–667, 2016.

Alexis B Webb, Iván M Lengyel, David J Jörg, Guillaume Valentin, Frank Jülicher, Luis G Morelli, and Andrew C Oates. Persistence, period and precision of autonomous cellular oscillators from the zebrafish segmentation clock. *eLife*, 5:e08438, 2016.

Zhaohua Wu and Norden E Huang. Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Advances in Adaptive Data Analysis*, 1(01):1–41, 2009.

Zhaohua Wu, Norden E Huang, and Xianyao Chen. The multi-dimensional ensemble empirical mode decomposition method. *Advances in Adaptive Data Analysis*, 1(03):339–372, 2009.

Zhaohua Wu, Jiaxin Feng, Fangli Qiao, and Zhe-Min Tan. Fast multidimensional ensemble empirical mode decomposition for the analysis of big spatiotemporal datasets. *Phil. Trans. R. Soc. A*, 374(2065):20150197, 2016.