Dynamics of positional information in the vertebrate neural tube

Anđela Marković[1,✉], James Briscoe[2], and Karen M Page[1,3]

[1]Department of Mathematics, University College London, London WC1E 6BT, United Kingdom
[2]The Francis Crick Institute, London NW1 1AT, United Kingdom
[3]Institute of Physics of Living Systems, University College London, London WC1E 6BT, United Kingdom

In developing embryos, cells acquire distinct identities depending on their position in a tissue. Secreted signaling molecules, known as morphogens, act as long-range cues to provide the spatial information that controls these cell fate decisions. In several tissues, both the level and the duration of morphogen signaling appear to be important for determining cell fates. This is the case in the forming vertebrate nervous system where antiparallel morphogen gradients pattern the dorsal-ventral axis by partitioning the tissue into sharply delineated domains of molecularly distinct neural progenitors. How information in the gradients is decoded to generate precisely positioned boundaries of gene expression remains an open question. Here, we adopt tools from information theory to quantify the positional information in the neural tube and investigate how temporal changes in signaling could influence positional precision. The results reveal that the use of signaling dynamics, as well as the signaling level, substantially increases the precision possible for the estimation of position from morphogen gradients. This analysis links the dynamics of opposing morphogen gradients with precise pattern formation and provides an explanation for why time is used to impart positional information.

positional information | neural tube | information theory | tissue development| morphogen

Correspondence: ***andela.markovic.16@ucl.ac.uk***

DRAFT

Introduction

The process of tissue development is remarkably reliable and precise. It involves the generation of distinct cell types in a characteristic and reproducible arrangement to form organised patterns of cell fates. In an influential essay that describes this process, Lewis Wolpert introduced the concept of "positional information" (1). This proposed that positional information is derived from signals, usually referred to as morphogens, which form concentration gradients in the tissue, with cells measuring the local concentration and interpreting it to select the appropriate cell fate for that position. In this view, concentration thresholds define boundaries that distinguish adjacent sets of cell types.

The dorsal-ventral patterning of the vertebrate nervous system is a well established system for studying morphogen-driven patterning (2). Signals emanating from two opposing sources – the dorsal roof plate secreting the morphogen BMP, and the ventral floor plate secreting the morphogen Shh – form antiparallel gradients that partition the neural tube into 11 discrete domains of molecularly distinct neural progenitors arrayed along the dorsal-ventral axis, Fig.1. Neural progenitors respond to morphogen signaling by controlling a gene regulatory network that specifies the pattern of gene expression along the dorsoventral axis and hence the positions at which distinct neuronal subtypes are generated. Gene expression and consequently cell fate appear to be specified by a combination of Shh and BMP signaling (3). Positions of several gene expression boundaries have been experimentally measured (3–6) and this indicates that the boundaries are specified with a high degree of precision, resulting in sharply defined transitions in gene expression and only moderate intermixing of cell types at domain boundaries.

To explore the amount of positional information that can be obtained from the two morphogens, we took advantage of tools from *information theory* (7). These have been used previously to study information transmission through signaling pathways and gene regulatory networks (8–16), including positional information transmission by the morphogen Bicoid, which governs the anterior-posterior patterning of the Drosophila blastoderm (17–21). The variability between embryos of the Bicoid gradient and the expression levels of the genes it controls was used to connect positional information to experimental data at a specific developmental time point. This revealed that Bicoid has a lower precision than its target genes (17), leading to the possibility that other sources of information or mechanisms are involved, such as continuous, rather than instantaneous assessment of Bicoid concentration (22).

The Drosophila blastoderm differs in several ways from many developing tissues. It is a syncitium. This allows Bicoid to regulate gene expression directly without an intervening signal transduction pathway. Most developing tissues, including the neural tube, are cellularised and morphogens spread extracellularly to activate intracellular signaling pathways that control gene expression (23, 24). In addition, the patterning of the blastoderm occurs in the absence of tissue growth (25). Most tissues grow as they are patterned, and morphogen gradients do not always scale with growth (6, 23). This, along with other factors such as changes in the production of the signal, results in the gradient changing as the tissue is patterned (23). Consequently, there is no straightforward correlation between the concentration of morphogens, their location in tissue, and cell fate allocation. Indeed, the duration of morphogen signalling has also been shown to influence cell fate decisions (4, 26–28). It is unclear, given the dynamics and the heterogeneity and fluctuations in the spread and interpretation of a morphogen, how a graded morphogen produces the precise patterns of cell fate decisions that are

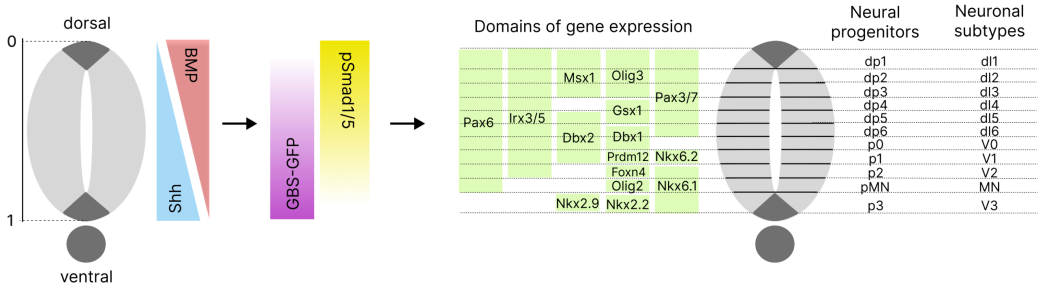

Fig. 1. The morphogen, Shh, emanating from the ventral pole, and the morphogen, BMP, secreted from the dorsal pole of the neural tube, form antiparallel gradients along the dorsal-ventral axis. These extracellular gradients are translated into intracellular gradients of the activity of Gli and Smad transcription factors, regulated by Shh and BMP signaling, respectively. Their activity is measured using GBS-GFP, a Gli reporter, and pSmad1 / 5, the activated version of Smad1/5. The Gli and Smad transcription factors control gene regulatory networks that specify the domains of gene expression along the axis, and hence, partition the neural tube into 11 discrete domains of molecularly distinct neural progenitors. In this way, the positions at which distinct neuronal subtypes will be generated are determined.

observed in developing tissues.

Despite the prominence of Wolpert's positional information concept, it has become evident that relying solely on morphogen concentration does not provide a comprehensive explanation for understanding morphogen gradients and their role in tissue development (24). The "progressive and dynamic manner" (26) in which the domains of neural progenitors appear suggests new ways of encoding positional information. Here, we set out to extend the information-theoretic approach by adapting methods that accommodate dynamics to analyse positional information in the neural tube. We wanted to explore the impact of signaling dynamics on positional precision and introduce a temporal dimension to the concept of positional information.

Results

Quantifying positional information with information theory. We first set out to quantify how much positional information Shh and BMP morphogens provide in the neural tube. The signaling pathways of these morphogens represent complex, noisy channels through which the cells receive information about their positions (23). Information theory describes the input and the output of any noisy channel as two random variables, X and Y, respectively (29). The noise in the channel leads to a distribution of possible output responses produced by the same input message (16), and the uncertainty in the input arises from the noise in its source as well as other features of the system. Let X be a discrete random variable that can take values $x_1, x_2, ..., x_m$, with the probability distribution $P_X(X) = (P(x_1), ..., P(x_m))$, and let Y be a continuous random variable, with the probability distribution $P_Y(Y)$ and $\mathcal{Y}$ as a space of all possible values of Y. The main measure of information that we use here, *mutual information* between two random variables X and Y, $MI(X,Y)$, is defined by (30)

$$MI(X,Y) = \sum_{i=1}^{m} \int_{\mathcal{Y}} P(x_i, y) \log_2 \frac{P(x_i, y)}{P_X(x_i) P_Y(y)} dy \quad \textbf{(1)}$$

where $P(x_i, y) = P(y|x_i)P_X(x_i) = P(x_i|y)P_Y(y)$. Mutual information $MI(X,Y)$ measures to what extent X and Y depend on each other, being 0 when they are independent, and captures any kind of correlation between the input and the output (19). It indicates how much the uncertainty about the value of the random variable X is reduced by knowing the state of Y, that is, how much information about the input we obtain after looking at the output (29). More precisely, $MI(X,Y)$ can be interpreted as the average number of binary (yes/no) questions about the value of X that are answered after observing Y. Each question divides possible values of X into two equally probable, mutually exclusive classes based on the values of Y associated with them. Hence, $2^{MI(X,Y)}$ gives us the average number of equiprobable classes of the values of the input X that produce classes of the values of the output Y that can be distinguished without error (31). Therefore, the maximum possible value of mutual information in this case is $\log_2 m$. Two individual values of X, x_i and x_j, are said to be resolved if their corresponding output distributions $P(Y|X = x_i)$ and $P(Y|X = x_j)$ do not overlap (32). When all the input values have the same probability, $P(x_1) = ... = P(x_m)$, then the mutual information depends solely on the overlap between their output distributions and provides a measure of the noise which leads to these overlaps and prevents the system from perfectly resolving each input value.

The outputs of the Shh and BMP signaling pathways are the activities of their intracellular transcriptional effectors. These were measured in the mouse neural tube by Zagorski et al. (3). Shh signaling results in the intracellular activation of Gli transcription factors, the activity of these was measured using a transcriptional reporter GBS-GFP. BMP signaling activates Smad transcription factors, which were measured using an antibody specific for phosphorylated SMAD1/5, pSmad1/5, the activated version of the BMP transcriptional effectors. Hence, we take the activities of GBS-GFP and pSmad1/5 to be the outputs of the Shh and BMP signaling pathways, respectively. To quantify the positional information provided by the signaling of Shh and BMP morphogens in the neural tube, we computed $MI(X,Y)$ where the input X represents different relative positions $x_1, ..., x_m$ along the dorsal-ventral length of the neural tube, and the output Y the levels of GBS-GFP and pSmad1/5 activity measured at these positions. Following the ideas in (17, 19), we argue that this definition of $MI(X,Y)$ is a proper measure of positional information in

DRAFT

the neural tube.

More precisely, the input probability distribution $P(X)$ represents the positions of cells along the dorsal-ventral axis. Since cell density is approximately constant along this axis, we take the distribution $P(X)$ to be uniform (6). The relative dorsal-ventral length of the neural tube from position 0.055 to 0.945, where the position 0 corresponds to the dorsal end and 1 to the ventral end, is divided into bins of width 0.05 with a distance of 0.01 between them. This resulted in 15 bins and therefore 15 input values denoted sequentially as $x_1, ..., x_{15}$ (*Supplementary Information* (SI) Fig.1). Using the measurements of the neural tube length over time from (3), the average cell diameter, 4.9μm (6), is 0.05 of the relative length at 5h of neural tube development, the first time point at which the signals are measured. The absolute length of each bin grows over time, but during the first ~ 30 hours of development, this length is less than 2 average cell diameters, and at 35h it is approximately 2 cells.

Zagorski et al. (3) measured the levels of GBS-GFP and pSmad1/5 every 0.01 of relative length from position 0.055 to 0.945, along the dorsal-ventral length of the neural tube in a single embryo. Hence, one of our 15 bins, x_i, comprises 6 single relative positions at which the signals are measured and the output values corresponding to x_i are concatenated measurements of the signals made at those 6 positions in all embryos.

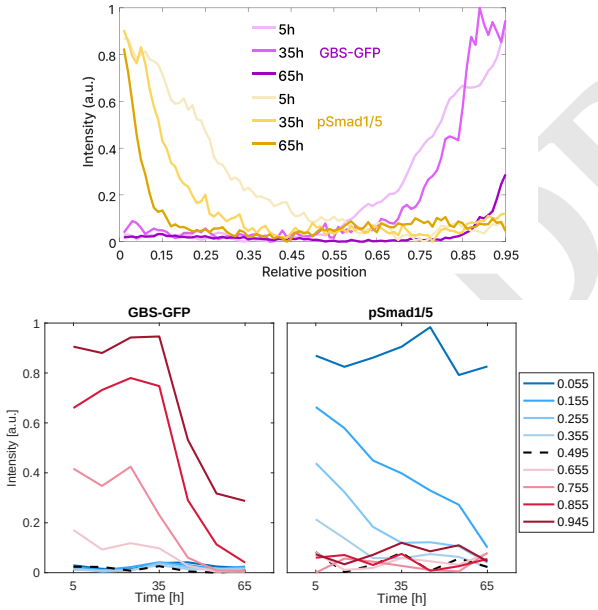

Fig. 2. Top: the mean gradients of the activity of GBS-GFP, the reporter of Shh signaling, and the gradients of the activity of pSmad1/5, the transcriptional effectors of the morphogen BMP, along the relative dorsal-ventral length of the neural tube at 5h, 35h and 65h during the development. Bottom: the mean levels of GBS-GFP and pSmad1/5 over time at the indicated relative positions.

Information-theoretic tools have been used before to study dynamic signals transmitting information to cells about their environment (33–35). This has led to the idea that a single cell monitors signals over time to overcome noise (33) and that information can be encoded in different dynamical patterns of signals (36). Therefore, the output values were represented as vectors of cellular responses at different time points. To explore the impact of morphogen signaling dynamics on positional precision, we took a similar approach and used the levels of GBS-GFP and pSmad1/5 measured at a particular relative position at successive time points, $t_1, t_2, ...$, during neural tube development. Hence, a single output response to an input x_i becomes a vector of the form $y^i = \left(y^i_G(t_1), y^i_G(t_2), ..., y^i_S(t_1), y^i_S(t_2), ...\right)$ where G denotes GBS-GFP and S pSmad1/5. Note that the signals are only measured once in an embryo and then the temporal trajectory is inferred from measurements at equivalent positions in embryos of sequential time classes (see *Supplementary Information* and (3, 6) for details).

To compute mutual information, we took advantage of a recently developed framework proposed by Jetka et al. (32) – SLEMI (statistical learning estimation of mutual information) – which simplifies the calculation of mutual information for high-dimensional output. The method involves a classifier that uses a logistic regression model to estimate $P(x_i|y)$, the probability of being at a position x_i after observing a signal level y, from the given data such that the probability of being at the true position for a given signal is maximised. Therefore, SLEMI belongs to a class of "decoding-based information estimators", as described in (37), which provide a lower bound to the information between X and Y (34). Jetka et al. (32) demonstrated the validity of their method and showed its advantages over the most commonly used method for computing mutual information, k-nearest neighbour.

The calculation of mutual information led to the values of $2^{MI(X,Y)}$ which, in this case, represent the average number of spatial regions along the dorsal-ventral axis of the neural tube that can be clearly delineated based on received signal levels. Since the positions are uniformly distributed, these regions are of equal length.

Decoding Morphogen Gradients at Single Time Points. We computed mutual information (MI) between the positions along the dorsal-ventral axis of the neural tube and the levels of GBS-GFP and pSmad1/5 activity measured at these positions. We considered the levels of each signal separately and the combined levels. The data consist of 102 measurements of each of the two signals measured in each of the 15 spatial bins at each of the 7 time points 5h, 15h, 25h, 35h, 45h, 55h and 65h during neural tube development. We first calculated the mutual information for each time point by using only the signal levels measured at the corresponding time point. We then considered series of signal levels at successive time points and computed the time-series mutual information from the first time point up to each later time point. For example, the values measured at 5h, 15h, 25h and 35h were used to compute the mutual information for the 35h time course.

The time-point mutual information is highest at the first three time points, 5h, 15h, 25h, Fig.3. However, when taking GBS-GFP and pSmad1/5 separately, the time-point MI does not reach 1 bit, the value needed to distinguish two regions. This indicates that at a single point in time, on average, the in-

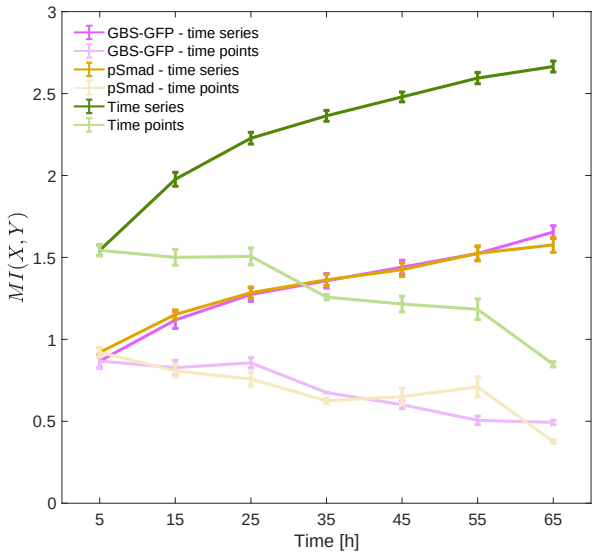

Fig. 3. Mutual information (MI) between the 15 positions along the dorsal-ventral axis and the levels of the signals GBS-GFP and pSmad1/5 measured at these positions. The values of MI are first computed using the levels of GBS-GFP and pSmad1/5 separately, and then based on the levels of the two signals combined (where it is not otherwise denoted). The levels were measured at 7 time points during development. The time-point mutual information was computed using signal levels measured at each time point individually. The time-series mutual information is computed using the series of signal levels from the first time point up to each later point. The values are averaged over different embryos at a single time point and over different combinations of embryos when time series is considered. The error bars are the standard deviations. Units are bits.

formation from either Shh or BMP signaling is insufficient to create a sharp boundary in the neural tube. On the other hand, the mutual information for the two signals combined is higher and shows that, at these first three time points, there is almost enough information to be able to resolve 3 regions ($2^{1.58} \approx 3$). Thus, by sensing both morphogens at a single time point, there can be sufficient information to specify 2 spatial regions.

Temporal Information from Morphogen Gradients Increases Positional Information. The time-series mutual information increases with the inclusion of an increasing number of time points, indicating that new information continues to accumulate over time. The information accumulated during the first 65h from only Shh signaling is sufficient to distinguish 3 regions along the neural tube. For pSmad1/5 activity, the MI values are similar. By contrast, the values of mutual information for the combined Shh and BMP signaling indicate that the information accumulated during the first ∼30h is sufficient to resolve 5 regions without error ($2^{2.32} \approx 5$). This rises to 2.67 ± 0.03 bits by 65h ($2^{2.58} \approx 6$). This confirms that the two morphogens acting together carry significantly more information. Overall, the analysis indicates that the combined signaling of the two morphogens over time changed the result from no two regions resolved at a single time point to almost 7 distinguishable regions accumulated over 65h.

We expected the time-series information to increase with time, the question is how quickly. The increase in the information becomes smaller at each next time step (SI Fig.5). However, using the interpretation of $2^{MI(X,Y)}$, the time-series mutual information for the combined signaling revealed that the average absolute length of the regions that can be perfectly distinguished remains approximately the same during the first ~ 30 hours, approximately 6 cell diameters. After that, the length increases; it is ~ 7 cells for the 35h time course and ~ 12 cells for the 65h of signaling. If we used the time-point mutual information, this length would be $6, 8, 10, 16, 20, 26, 42$ at the seven time points, respectively. This means that during the first ~ 30h of development, the average amount of information that is accumulated from Shh and BMP signaling increases proportionally to the tissue growth, maintaining positional precision. This result is in agreement with the claim that the pattern is established during the first ~ 30h, which was supported by the analysis of Zagorski et al. (3) and the experimental evidence of Kicheva et al. (6).

This result is dependent on the choice of the time points and this level of precision is lost when we consider the points every 20h, 30h or 60h (SI Fig.4). As we increase the time step, the total amount of information for the 65h time course significantly decreases, while the information gain at the first/second/third time step, in general, becomes slightly lower (SI Fig.5). Here, we assume that the signal levels received over time are stored. This comes with a cost, which increases with the dimensionality of the time series, which is, on the other hand, limited by the memory capacity of the cells (35). Another disadvantage of longer time series is that, since the time points are closer, there will be more redundant information (35). In our case, considering all the 7 time points leads to a high level of redundant information, close to the maximal values (SI Fig.6), which raises the question of the efficiency of encoding information in this way.

Comparison of Positional Information from Time Series and Time Integrals. The dynamics of neural tube patterning challenges both the concept of positional information and the information-theoretic approach to studying dynamic signals. If the duration during which signal levels are maintained above a particular threshold determines position (26), the question is whether a whole vector of signaling dynamics should represent encoded positional information. It has been suggested that positional information in the neural tube could be encoded as the time integral of signaling (26, 28). Therefore, we considered cumulative level and duration by summing the entries of vectors for each of the two signals and computed $MI(X,Y)$ where a single output response corresponding to an input x_i is now $y^i = \left(y^i_G(t_1) + y^i_G(t_2) + ..., y^i_S(t_1) + y^i_S(t_2) + ...\right)$. This time-integral MI during the first ~ 30h is slightly lower than the time-series mutual information, Fig.4. The difference between the two slowly increases over time since the time-integral mutual information plateaus after 35h, with even a slight decrease at 65h. If the domains of gene expression are established during the first ~ 30h, any new information that arrived after that time would not be used. Furthermore, when considering the two signals together, the time-integral mutual information gives approximately the same level of precision during the first ~ 30h as the time-series mutual information,

DRAFT

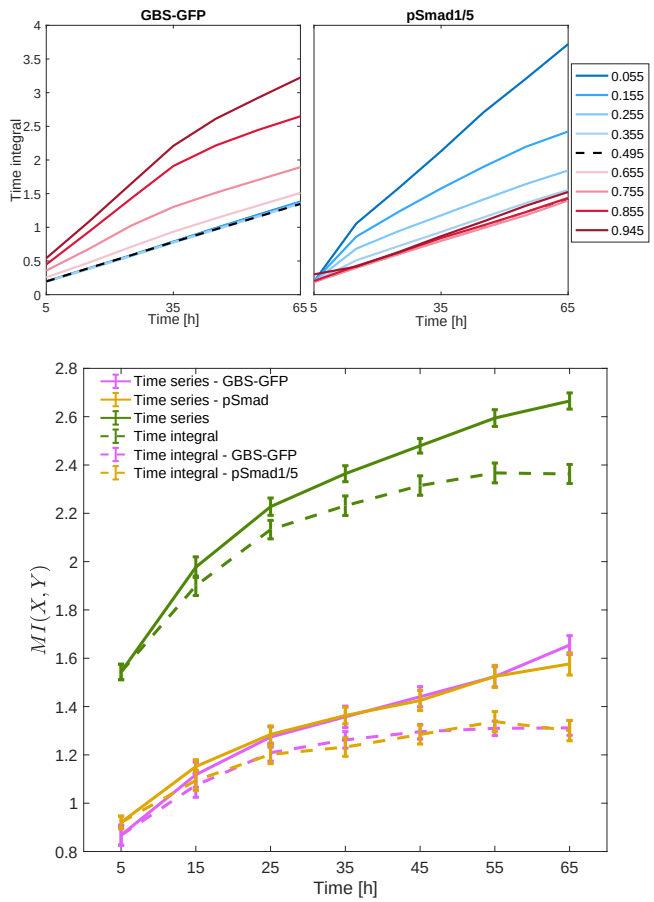

Fig. 4. Top: the mean values of time integrals of GBS-GFP and pSmad1/5 over time at the indicated relative positions. Bottom: the time-series and the time-integral mutual information using the levels of GBS-GFP and pSmad1/5 separately, and based on the levels of the two signals combined. The time-integral mutual information is measured by summing the signal levels received over time, up to each time point. Units are bits.

i.e. the absolute length of the regions along the neural tube that can be perfectly distinguished is ~ 6 cells during this time. Later, the length increases and is around 15 cells by 65h.

Encoding information using combined dynamics of multiple factors has also been considered by Granados et al. (34). They introduced a measure for the degree of redundancy between a pair of transcription factors which gives 0 when they are completely independent and 0.5 when they are completely redundant. In the time-series scenario, the redundancy between the activities of GBS-GFP and pSmad1/5 is approximately 0.13 during the first 45h and then increases to 0.18 by 65h. When considering time integrals of the signals, the redundancy decreases over time to less than 0.1 by 65h (SI Fig.7).

These results support experimental findings (26) which suggested that the time integral is a suitable mechanism for encoding positional information in the neural tube. In any case, the introduction of time led to considerably more information than if static morphogen gradients were considered. This may explain why developmental systems respond to duration of morphogen signaling as well as its level. It emphasises the importance of a definition of positional information that captures this extended morphogen sensing.

Probability Distributions of Positional Assignments. The number of bits does not necessary indicate the number of regions that are perfectly distinguishable in the neural tube: 1 bit could indicate that two regions are completely resolved or it could indicate that 3 or more regions are resolved, but with some error (38). As mentioned above, SLEMI tells us how well the true position of an observed signal level can be inferred by estimating the probabilities $P(X = x_j|Y = y_l^i)$, for all $i, j = 1, ..., 15$ and $l = 1, ..., 102$, where y_l^i denotes l^{th} output response measured in bin x_i. Here, a signal level y_l^i can be either a scalar, when the level of GBS-GFP or pSmad1/5 is taken separately at a single time point or its time integral, or a vector when both signals and/or consecutive time points are considered. Since signal levels are measured at 90 relative positions, every 0.01 of relative length from position 0.055 to 0.945, we find the average probability of each of these positions being in any of the 15 bins by averaging $P(X = x_j|Y = y_l^i)$ over signal levels measured at that position (21). We denote these probabilities by $P(X^* = x_j|x^{i,k})$ where X^* is a discrete random variable with values x_1,..., x_{15} corresponding to the bins and $x^{i,k}$, for $i = 1, ..., 15$ and $k = 1, ..., 6$, denotes one of 6 relative positions in a bin x_i.

Using the levels of GBS-GFP and pSmad1/5 activity together at individual time points, the probabilities of classifying relative positions into correct bins along the patterning axis increase at the poles of the neural tube and decrease everywhere else with each next time point (Fig.5, SI Fig.8). The lack of scaling of the gradients as the neural tube grows leads to this expansion of the region of low positional information over time (Fig.2). On the other hand, the time-series analysis shows that information accumulated during the first 35h leads to notably higher probabilities of correct positional assignment along the whole length of the neural tube, particularly for the positions within the dorsal and ventral third. Considering the time integral of signal levels, the information along the neural tube is distributed in the same way as in the time-series scenario. The time integrals of signal levels lead to slightly lower probabilities during the first 35h, compared to the time-series scenario, and this difference increases over time. More precisely, the highest difference between the time-series and time-integral probabilities after 15h, 25h and 35h is 0.06, 0.09 and 0.09, respectively, while after 65h it is 0.27 (SI Fig.9).

We looked closely at the average distribution $P(X^*|x^i)$, the probabilities of being in any bin after observing a signal level measured at a single relative position x^i within bin x_i, for $i = 4, 8, 12$ and $x^4 = 0.255, x^8 = 0.495$ and $x^{12} = 0.755$, at time point 5h and for time series and time integral over the first 25h, Fig.6. The increase in the probability of being in the correct bin between 5h and 25h is particularly noticeable for the positions near each pole of the neural tube, where one signal is maintained above a certain threshold during the first 25h, while the levels of the other signal take the lowest values (Fig.2). This is consistent with the findings that the duration of signaling influences cell fate decisions. For the

DRAFT

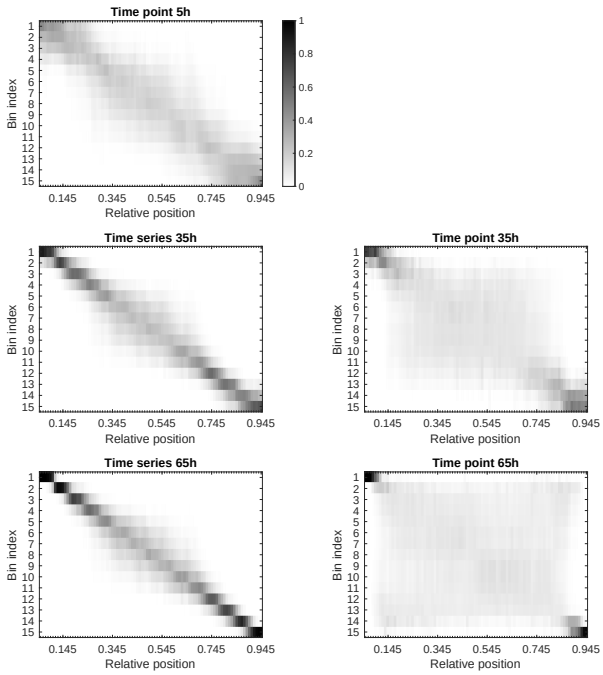

Fig. 5. The average probability of being in any of the 15 bins after observing levels of the two signals measured at one of the relative positions from 0.055 to 0.945, at time points 5h, 35h and 65h, and for the time series of levels up to 35h and 65h.

middle of the neural tube, where both signals take the lowest values, information also accumulates over time and the probabilities of correct discrimination increase, although to a notably lesser extent resulting in the less marked increase in the peak of the distribution.

Observing the time-series probabilities for the 35h time course obtained from the levels of GBS-GFP and pSmad1/5 activity separately, it is evident that the signaling from the second morphogen contributes significantly to precision along the major part of the neural tube, Fig.7. Shh signaling alone provides insufficient information for the region $[0.055, 0.585]$ of the neural tube, and BMP signaling provides limited information to $[0.445, 0.945]$, where the probabilities of correct discrimination are around 0.1.

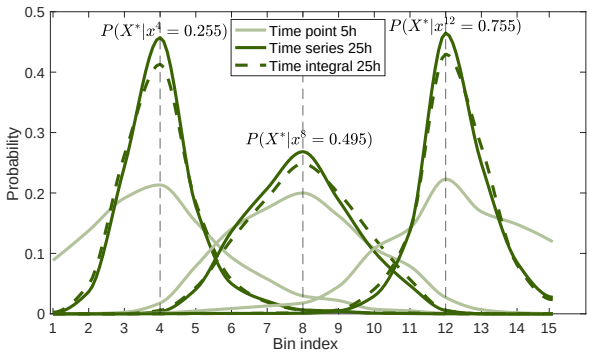

Fig. 6. The probability distributions $P(X^*|x^i)$ showing the probabilities of being in any bin after observing a signal level measured at a relative position x^i, for $x^4 = 0.255, x^8 = 0.495$ and $x^{12} = 0.755$. We computed these at time point 5h and for time series and time integral over the first 25h .

Spatial Patterns of Positional Uncertainty. A signal level y_l^i is correctly decoded if the probability distribution $P(X|Y = y_l^i)$ attains its maximum at the correct bin x_i, and incorrectly if the maximum is at any other x_j, $j \neq i$. We assume that this maximum is always unique since the system should provide unambiguous positional information. Therefore, the fraction of incorrectly decoded signal levels at time points 5h, 15h and 25h is 32% , 30% and 34%, while for the 15h, 25h, 35h and 65h time courses it is 16%, 12% , 11% and 5%, respectively. When we consider the time integrals after the first 15h, 25h and 35h, this fraction is 20%, 17% and 19%. Examining the relative positions that are incorrectly decoded at the time point 5h indicated that these positions are almost uniformly distributed along the axis, with the most incorrectly discriminated positions at the ventral end, Fig.8. Taking the time series and the time integral after the first 25h decreases the number of incorrectly decoded levels in the ventral and dorsal third of the neural tube, while in the middle only the time series results in a significant improvement.

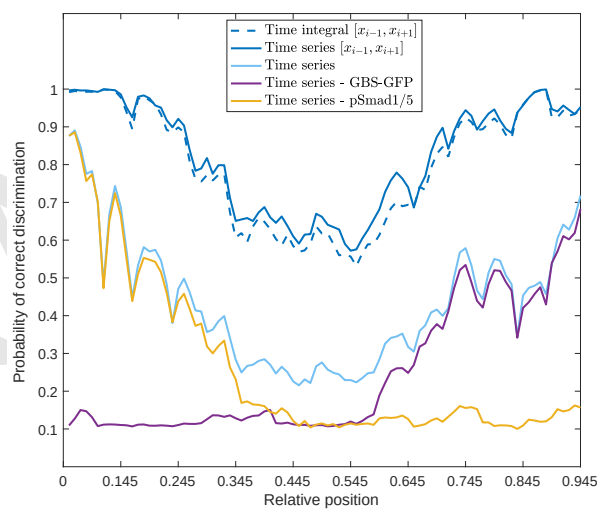

Fig. 7. The average probabilities of each of the relative positions being correctly classified in its own bin (where it is not otherwise denoted) or in the region containing the correct bin and its adjacent bins, denoted by $[x_{i-1}, x_{i+1}]$. All the probabilities are computed based on the levels of the signals during the first 35h, using the time series or the time integral. The probabilities are obtained based on the two signals together and, in the time-series scenario only, using the levels of the two signals separately as well.

The time-series and time-integral results showed that after 25h and 35h a correct bin x_i is mistaken for a bin adjacent to it, i.e. for x_{i-1} and x_{i+1} (or only x_{i+1} if $i = 1$ and x_{i-1} if $i = 15$). In terms of relative length, this is within the region ± 0.06 from the correct bin, which is ~ 2.5 cell diameters at 35h. On the other hand, at the first three time points individually, 10%, 15%, 21% of incorrectly decoded signal levels, respectively, are most probable to be in bins that are not adjacent to the correct one.

We calculated the average probability of being inside an area including x_{i-1}, x_i and x_{i+1} – the true bin and its nearest neighbours – after observing a signal

level measured at each of the considered relative positions, found as the sum $P(X^* = x_{i-1}|x^{i,k}) + P(X^* = x_i|x^{i,k}) + P(X^* = x_{i+1}|x^{i,k})$. We denote this sum as $P(X^* \in [x_{i-1}, x_{i+1}]|x^{i,k})$. Compared to the average probabilities of being in the correct bin, the average probabilities of being in the region $[x_{i-1}, x_{i+1}]$ take considerably higher values. In particular, considering the time integral and the time series over the first 35h, the average probabilities of being in the region $[x_{i-1}, x_{i+1}]$ are higher than 0.5 for all positions, and for the dorsal region $[0.055, 0.255]$ and the ventral region $[0.695, 0.945]$ these probabilities are higher than or equal to 0.9, Fig.7. If we assume that $P(X^*|X = x^{i,k})$ is a normal distribution, then $\overline{P}(X^* \in [x_{i-1}, x_{i+1}]|x_i)$ and $[x_{i-1}, x_{i+1}]$ as a confidence interval can be used to approximate its standard deviation. The mean is the bin x_i, whose absolute length grows from ~ 1 to ~ 2 cell diameters between 5h and 35h. At the first four time points individually, the highest errors are found in the middle of the neural tube and are equal to $\sim 2, 3, 4$ and 7 cell diameters, respectively. At the same points, in both the time-series and time-integral scenarios, the errors are $\sim 2, 2, 2$ and 3 cells, respectively. This is consistent with the idea that signaling dynamics improves precision and maintains it at approximately the same level during the first $\sim$ 30h.

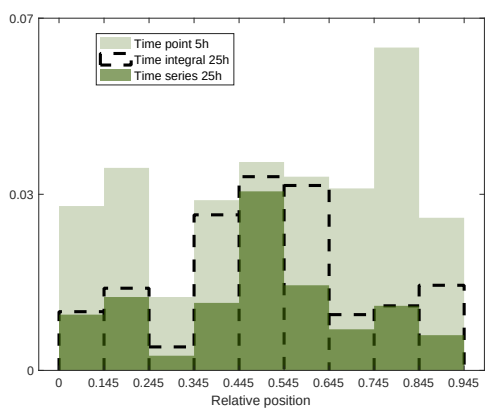

Fig. 8. The distributions of incorrectly decoded signal levels at time point 5h and after the first 25h, using the time integral and the time series . The x-axis shows the relative positions where the incorrectly decoded signal levels are measured and y-axis the fraction they make out of all the signal levels considered at the given time. The results are obtained based on the two signals together.

Discussion

The development of the neural tube requires precise spatial patterning along the dorsal-ventral axis. This is governed by Shh and BMP morphogen gradients that provide positional information via intracellular signaling pathways (25). Using experimentally measured levels (3) of the downstream transcriptional effectors of these signaling pathways – GBS-GFP and pSmad1/5 activity, respectively – we used the SLEMI algorithm (32) to quantify and decode the positional information transmitted by the two pathways. This analysis confirmed that a single morphogen gradient contains insufficient information to accurately resolve positions, except in the most dorsal or ventral regions of the neural tube, adjacent to the source of morphogen. By contrast, the combination of Shh and BMP signaling provides substantially more spatial information. Furthermore, we found that both the temporal trajectories and the time integrals of the signals carried more information than instantaneous snapshots. Both methods of encoding information using temporal profiles provide very similar results during the first 35 hours, while later only time series contribute new information. However, in both cases, during the first 35 hours there is sufficient information for accurate positional estimates, particularly for the dorsal and ventral thirds of the neural tube. Thus, cells would gain significantly more information by tracking signaling over the initial 35 hours of neural tube development and this would allow a substantially increased precision of positional decoding in the neural tube. Nevertheless, in the middle third of the neural tube, where both BMP and Shh signaling is low, the information provided by the measured signaling gradients remained limited. Overall, therefore, tracking the temporal profile of the two signals improved precision, and the analysis is consistent with the temporal profile of the two signals together providing more positional information than either signal individually or snapshots of the signals.

Previously Zagorski et al. (3) established a mathematical framework for decoding positional information in the neural tube. This involved several approximations. The probability distribution of signal at a single relative position x, $P(Y|X = x)$, was assumed to be Gaussian. The signals were also assumed to be independent and only individual time points considered. Our approach relaxes these assumptions. Nevertheless, the conclusions are broadly consistent with the previous analysis. Similar to the results reported here, Zagorski et al. show that the time-point positional error is highest for the middle of the neural tube and increases over time; they found the highest positional imprecision to be $\sim 3, 4, 5, 8$ cells at 5h, 15h, 25h, 35h, respectively. Moreover, Zagorski et al. conclude that a combination of signaling provides substantially lower positional error throughout the neural tube than either signal individually. The analyses are consistent with previous experimental evidence that the initial 30h of mouse neural tube development are critical for tissue patterning. Our approach extends the analysis from Zagorski et al. by taking series and integrals of signal levels during the initial $\sim$ 30h to conclude that the signaling over time decreases positional error. Patterning specifies 11 domains of neural progenitors. Assuming that these domains are each of the same size, the length of each would be ~ 3 cell diameters at $\sim$ 30h. We found that, at this point in patterning, the accumulated information is sufficient to perfectly distinguish regions containing ~ 6 cells on average. Moreover, the highest error, given by the standard deviation, with which a bin of size less than 2 cell diameters can be specified is ~ 2 cells. Therefore, a domain containing 3 cells could be specified with an error of less than 2 cells. Zagorski et al. have measured the boundaries of the two domains of gene expression, Pax6 and Nkx6.1, which are positioned in the middle of the tube. The precision of the boundaries was found to

be approximately ± 1 and ± 2 cells at 30h. As suggested by Zagorski et al. (3) , the similarity in precision of signaling profiles and boundary positions indicates that cells decode the information in a way that minimises imprecision.

For optimal positional information transmission, the precision should be constant along the entire neural tube length (17). However, our results show the discrepancy in precision between the middle and the ends of the neural tube, which raises the question if this is potentially compensated for by another source of information or another mechanism for information decoding. The gene regulatory networks downstream of morphogen signaling play a crucial role in the interpretation of positional information (2, 25). It is still unknown how precisely the levels of Shh and BMP signaling are measured downstream. If only thresholds of their concentrations can be sensed (39), the question is how much positional information is used. However, cross-repressive interactions between transcription factors help refine boundaries, as evidenced by shifting, noisy boundaries when specific transcription factors in the downstream network are lost (5). The design of the network appears to enhance patterning precision by reducing the effects of transient fluctuations in signaling. Extending the information-theoretic approach directly to quantified transcription factor levels could provide further insight into how the gene regulatory network processes time-varying positional information and affects the precision of tissue patterning. In addition, corrective mechanisms such as differential adhesion have been implicated in contributing to patterning precision (24). Assessing the role such mechanisms play in the precision of neural tube patterning would provide insight into the relative contribution of the information provided by morphogen signaling compared to downstream mechanisms.

The information-theoretic approach provides a powerful framework for analysing morphogen-based patterning that does not rely on detailed mechanistic knowledge of the patterning process. The data-driven nature, however, means that the accuracy of the experimental measurements of signaling dynamics influences the analysis, and the analysis is constrained to the available time points at which measurements are made. In regions of the tissue where the level of signaling is low, the fluorescent intensity of the corresponding signal reporters is low and approaches background levels; this increases the uncertainty in estimated signaling levels and results in lower estimates of positional information than would be obtained if more sensitive reporters were available. In addition the temporal trajectories used in our measurements contain variations at the level of the whole population. It would be ideal to follow the level of signaling in individual cells in the same embryo over time. This would provide additional insight into how much information cells receive and allow quantification of the noise that arises in a single embryo.

Overall, the study supports the conclusion that information theory combined with quantitative dynamic data is a powerful approach to unravel the mechanisms of precise spatial patterning in development, and permits a quantifiable definition of *positional information*. The temporal integration of morphogen signals is likely crucial for the patterning of many developmental systems, as the duration of morphogen signaling has been implicated in the patterning of other tissues (40–42). The methodology described in this study could therefore provide insight into other developmental systems. The success will depend on obtaining the appropriate quantitative data and expanding theoretical and experimental knowledge of the systems.

ACKNOWLEDGEMENTS

We are grateful to Anna Kicheva and Marcin Zagórski for helpful comments and advice. AM's work was supported by a teaching assistantship from the Department of Mathematics, UCL. Work in JB's group is supported by the Francis Crick Institute, which receives its core funding from Cancer Research UK (CC001051), the UK Medical Research Council (CC001051), and the Wellcome Trust (CC001051); and by the Wellcome Trust (220379/D/20/Z).

Bibliography

1. L. Wolpert. Positional information and the spatial pattern of cellular differentiation. *Journal of Theoretical Biology*, 25(1):1–47, 1969. ISSN 0022-5193. doi: https://doi.org/10.1016/S0022-5193(69)80016-0.
2. Andreas Sagner and James Briscoe. Establishing neuronal diversity in the spinal cord: a time and a place. *Development*, 146(22):dev182154, 11 2019. ISSN 0950-1991. doi: 10.1242/dev.182154.
3. Marcin Zagorski, Yoji Tabata, Nathalie Brandenberg, Matthias P. Lutolf, Gašper Tkačik, Tobias Bollenbach, James Briscoe, and Anna Kicheva. Decoding of position in the developing neural tube from antiparallel morphogen gradients. *Science*, 356(6345):1379–1383, 2017. doi: 10.1126/science.aam5887.
4. N Balaskas, A Ribeiro, J Panovska, E Dessaud, N Sasai, KM Page, J Briscoe, and V Ribes. Gene regulatory logic for reading the sonic hedgehog signaling gradient in the vertebrate neural tube. *Cell*, 148(1-2):273–284, 2012.
5. Katherine Exelby, Edgar Herrera-Delgado, Lorena Garcia Perez, Ruben Perez-Carrasco, Andreas Sagner, Vicki Metzis, Peter Sollich, and James Briscoe. Precision of tissue patterning is controlled by dynamical properties of gene regulatory networks. *Development*, 148(4):dev197566, 02 2021. ISSN 0950-1991. doi: 10.1242/dev.197566.
6. Anna Kicheva, Tobias Bollenbach, Ana Ribeiro, Helena Pérez Valle, Robin Lovell-Badge, Vasso Episkopou, and James Briscoe. Coordination of progenitor specification and growth in mouse and chick spinal cord. *Science*, 345(6204):1254927, 2014. doi: 10.1126/science.1254927.
7. C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, 1948. doi: 10.1002/j.1538-7305.1948.tb01338.x.
8. Etay Ziv, Ilya Nemenman, and Chris H. Wiggins. Optimal signal processing in small stochastic biochemical networks. *PLOS ONE*, 2(10):1–16, 10 2007. doi: 10.1371/journal.pone.0001077.
9. Filipe Tostevin and Pieter Rein ten Wolde. Mutual information between input and output trajectories of biochemical networks. *Phys. Rev. Lett.*, 102:218101, May 2009. doi: 10.1103/PhysRevLett.102.218101.
10. Gašper Tkačik and William Bialek. Information processing in living systems. 2014.
11. Siobhan S. Mc Mahon, Aaron Sim, Sarah Filippi, Robert Johnson, Juliane Liepe, Dominic Smith, and Michael P.H. Stumpf. Information theory and signal transduction systems: From molecular information processing to network inference. *Seminars in Cell Developmental Biology*, 35:98–108, 2014. ISSN 1084-9521. doi: https://doi.org/10.1016/j.semcdb.2014.06.011. Regulated Necrosis Modeling developmental signaling pathways Development of connective maps in the brain.
12. Pankaj Mehta, Sidhartha Goyal, Tao Long, Bonnie L. Bassler, and Ned S. Wingreen. Information processing and signal integration in bacterial quorum sensing. *Molecular Systems Biology*, 5:325 – 325, 2009.
13. Christian Waltermann and Edda Klipp. Information theory based approaches to cellular signaling. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1810(10):924–932, 2011. ISSN 0304-4165. doi: https://doi.org/10.1016/j.bbagen.2011.07.009. Systems Biology of Microorganisms.
14. Alex Rhee, Raymond Cheong, and Andre Levchenko. The application of information theory to biochemical signaling systems. *Physical Biology*, 9(4):045011, aug 2012. doi: 10.1088/1478-3975/9/4/045011.
15. Matthew D Brennan, Raymond Cheong, and Andre Levchenko. How information theory handles cell signaling and uncertainty. *Science (American Association for the Advancement of Science)*, 338(6105):334–335, 2012.
16. Raymond Cheong, Alex Rhee, Chiaochun Joanne Wang, Ilya Nemenman, and Andre Levchenko. Information transduction capacity of noisy biochemical signaling networks. *Science*, 334(6054):354–358, 2011. doi: 10.1126/science.1204553.
17. Julien O. Dubuis, Gašper Tkačik, Eric F. Wieschaus, Thomas Gregor, and William Bialek. Positional information, in bits. *Proceedings of the National Academy of Sciences*, 110(41):16301–16308, 2013. doi: 10.1073/pnas.1315642110.
18. Gašper Tkačik and Aleksandra M Walczak. Information transmission in genetic regulatory networks: a review. *Journal of Physics: Condensed Matter*, 23(15):153102, apr 2011. doi: 10.1088/0953-8984/23/15/153102.
19. Gašper Tkačik and Thomas Gregor. The many bits of positional information. *Development*, 148(2):dev176065, 02 2021. ISSN 0950-1991. doi: 10.1242/dev.176065.

DRAFT

20. Thomas R Sokolowski, Thomas Gregor, William Bialek, and Gašper Tkačik. Deriving a genetic regulatory network from an optimization principle. 2023.
21. Mariela D. Petkova, Gašper Tkačik, William Bialek, Eric F. Wieschaus, and Thomas Gregor. Optimal decoding of cellular identities in a genetic network. *Cell*, 176(4):844–855.e15, 2019. ISSN 0092-8674. doi: https://doi.org/10.1016/j.cell.2019.01.007.
22. Tanguy Lucas, Huy Tran, Carmina Angelica Perez Romero, Aurélien Guillou, Cécile Fradin, Mathieu Coppey, Aleksandra M. Walczak, and Nathalie Dostatni. 3 minutes to precisely measure morphogen concentration. *PLOS Genetics*, 14(10):1–25, 10 2018. doi: 10.1371/journal.pgen.1007676.
23. Michael Cohen, Anna Kicheva, Ana Ribeiro, Robert Blassberg, Karen Page, Chris Barnes, and James Briscoe. Ptch1 and gli regulate shh signalling dynamics via multiple mechanisms. *Nature communications*, 6:6709, 04 2015. doi: 10.1038/ncomms7709.
24. Anna Kicheva and James Briscoe. Control of tissue development by morphogens. *Annual Review of Cell and Developmental Biology*, 39(1):91–121, 2023. doi: 10.1146/annurev-cellbio-020823-011522. PMID: 37418774.
25. James Briscoe and Stephen Small. Morphogen rules: design principles of gradient-mediated embryo patterning. *Development*, 142(23):3996–4009, 12 2015. ISSN 0950-1991. doi: 10.1242/dev.129452.
26. Eric Dessaud, Vanessa Ribes, Nikolaos Balaskas, Lin Lin Yang, Alessandra Pierani, Anna Kicheva, Bennett G. Novitch, James Briscoe, and Noriaki Sasai. Dynamic assignment and maintenance of positional identity in the ventral neural tube by the morphogen sonic hedgehog. *PLOS Biology*, 8(6):1–14, 06 2010. doi: 10.1371/journal.pbio.1000382.
27. Eric Dessaud, Lin Lin Yang, Katy Hill, Barny Cox, Fausto Ulloa, Ana Ribeiro, Anita Mynett, Bennett G Novitch, and James Briscoe. Interpretation of the sonic hedgehog morphogen gradient by a temporal adaptation mechanism. *Nature*, 450(7170):717–720, November 2007.
28. Samuel Tozer, Gwenvael Le Dréau, Elisa Marti, and James Briscoe. Temporal control of BMP signalling determines neuronal subtype identity in the dorsal neural tube. *Development*, 140(7):1467–1474, 04 2013. ISSN 0950-1991. doi: 10.1242/dev.090118.
29. T.M. Cover and J.A. Thomas. *Elements of Information Theory*. Wiley, 2012. ISBN 9781118585771.
30. Gašper Tkačik. From statistical mechanics to information theory: understanding biophysical information-processing systems, 2010.
31. Gašper Tkačik. Information flow in biological networks. 2007.
32. Tomasz Jetka, Karol Nienałtowski, Tomasz Winarski, Sławomir Błoński, and Michał Komorowski. Information-theoretic analysis of multivariate single-cell signaling responses. *PLOS Computational Biology*, 15(7):1–23, 07 2019. doi: 10.1371/journal.pcbi.1007132.
33. Jangir Selimkhanov, Brooks Taylor, Jason Yao, Anna Pilko, John Albeck, Alexander Hoffmann, Lev Tsimring, and Roy Wollman. Accurate information transmission through dynamic biochemical signaling networks. *Science*, 346(6215):1370–1373, 2014. doi: 10.1126/science.1254933.
34. Alejandro A. Granados, Julian M. J. Pietsch, Sarah A. Cepeda-Humerez, Iseabail L. Farquhar, Gašper Tkačik, and Peter S. Swain. Distributed and dynamic intracellular organization of extracellular information. *Proceedings of the National Academy of Sciences*, 115 (23):6088–6093, 2018. doi: 10.1073/pnas.1716659115.
35. Garrett D. Potter, Tommy A. Byrd, Andrew Mugler, and Bo Sun. Dynamic sampling and information encoding in biochemical networks. *Biophysical Journal*, 112(4):795–804, 2017. ISSN 0006-3495. doi: https://doi.org/10.1016/j.bpj.2016.12.045.
36. Lauritz Hahn, Aleksandra M. Walczak, and Thierry Mora. Dynamical information synergy in biochemical signaling networks. *Phys. Rev. Lett.*, 131:128401, Sep 2023. doi: 10.1103/PhysRevLett.131.128401.
37. Sarah Anhala Cepeda-Humerez, Jakob Ruess, and Gašper Tkačik. Estimating information in time-varying signals. *PLOS Computational Biology*, 15(9):1–33, 09 2019. doi: 10.1371/journal.pcbi.1007290.
38. Anders S Hansen and Erin K O'Shea. Limits on information transduction through amplitude and frequency regulation of transcription factor activity. *eLife*, 4:e06559, may 2015. ISSN 2050-084X. doi: 10.7554/eLife.06559.
39. Marianne Bauer, Mariela D. Petkova, Thomas Gregor, Eric F. Wieschaus, and William Bialek. Trading bits in the readout from a genetic network. *Proceedings of the National Academy of Sciences*, 118(46):e2109011118, 2021. doi: 10.1073/pnas.2109011118.
40. Engda G Hagos and Scott T Dougan. Time-dependent patterning of the mesoderm and endoderm by nodal signals in zebrafish. *BMC Developmental Biology*, 7(1):22, March 2007.
41. Keisuke Sako, Saurabh J. Pradhan, Vanessa Barone, Álvaro Inglés-Prieto, Patrick Müller, Verena Ruprecht, Daniel Čapek, Sanjeev Galande, Harald Janovjak, and Carl-Philipp Heisenberg. Optogenetic control of nodal signaling reveals a temporal pattern of nodal signaling regulating cell fate specification during gastrulation. *Cell Reports*, 16(3):866–877, 2016. ISSN 2211-1247. doi: https://doi.org/10.1016/j.celrep.2016.06.036.
42. Brian D Harfe, Paul J Scherz, Sahar Nissim, Hua Tian, Andrew P McMahon, and Clifford J Tabin. Evidence for an expansion-based temporal shh gradient in specifying vertebrate digit identities. *Cell*, 118(4):517–528, August 2004.

DRAFT