

Accuracy of positional information provided by multiple morphogen gradients with correlated noise

Yoshihiro Morishita^{1,2} and Yoh Iwasa¹¹*Department of Biology, Faculty of Sciences, Kyushu University, Fukuoka 812-8581, Japan*²*PRESTO, Japan Science and Technology Agency, 4-1-8 Honcho Kawaguchi, Saitama, Japan*

(Received 17 October 2008; revised manuscript received 5 February 2009; published 4 June 2009)

Normal development of multicellular organisms requires cells to respond properly according to their positions. Positional information is often provided to cells as concentrations of diffusive chemicals called morphogens with spatial gradients. However, the spatial profiles of their concentrations include various kinds of noises, making positional information unreliable. In many developmental systems, multiple morphogen gradients are adopted to specify the spatial position along a single axis, presumably to achieve a sufficiently high precision of information on the location of each cell. In this paper, we ask how the precision of positional information depends on the number of morphogens. We derive a formula for the limit of precision when each cell adopts the maximum-likelihood estimation of the “true” position from noisy inputs. The precision increases with the number of morphogens and interestingly it also depends on the correlation of noises. The positional specification can be made more precisely if their gradients are of the opposite (same) direction when noises of the two morphogens are positively (negatively) correlated. The formula also tells us a minimum number of morphogens needed to achieve a given precision of positional information. We illustrate the theory by analyzing experimental data for the gradients of two diffusive chemicals, Bicoid and Caudal, in the early development of *Drosophila* embryo. The analysis suggests that combined information provided by the two chemicals is able to give accurate positional information in the middle part of the embryo, where the embryo segmentation occurs in later stages, much more than near both ends.

DOI: [10.1103/PhysRevE.79.061905](https://doi.org/10.1103/PhysRevE.79.061905)

PACS number(s): 87.18.Tt, 87.17.Pq, 05.40.-a

I. INTRODUCTION

Normal development of multicellular organisms requires cells to grow, differentiate, divide, and die properly according to their spatial position [1]. During morphogenesis, concentration gradients of diffusive chemicals called morphogens are often established in different ways [2–7] and give cells clues on their locations [8–10]. For example, in a one-dimensional space, if a morphogen is secreted from one end and forms a gradient through diffusion and degradation, cells can recognize their distance from the end simply by sensing its concentration [10–12] [see Fig. 1(a)].

Developmental processes are performed under many kinds of noises. The body size and the reaction rates of synthesis and degradation of morphogens may be different among embryos [13–16]. Further, within each embryo, there may be noise in the diffusion process due to the spatial disorder and the finiteness of molecule number [17–20].

These noises make the positional information provided by morphogen concentrations ambiguous. Considering the importance of accurate specification of cell position, organisms are likely to have evolved mechanisms that allow performing robust positional specification despite the presence of noises. Possible noise-reducing mechanisms in the formation and interpretation of morphogen gradients have been discussed [21–27].

In the present paper, we consider the robustness of positional specification from a perspective different from the previous studies. To be specific, we focus on the number of morphogen species adopted in the positional specification. In a deterministic situation (i.e., in the absence of noises), the minimum number of morphogens needed to specify the po-

sition is equal to the spatial dimension. However, in many developmental systems, more chemical species than the spatial dimension are adopted [Fig. 1(b)]. For example, in the segmentation of *Drosophila* embryo, several diffusive chemicals including Bicoid, Caudal, and Nanos are used to provide positional information along a single anteroposterior axis [28–30]. Is positioning by more chemicals always more robust to noises? If so, how much is the precision improved with the number? To answer these questions is the main purpose of this study.

We first systematically examine the precision of the positional information given by multiple morphogens. To make the argument clear, we focus on the positional specification in a one-dimensional space and derive a formula for the limit of the precision of positional information under a given number of morphogens in the presence of noises. Instead of focusing on a particular intracellular biochemical dynamics to read out their gradients which have been discussed previously (see [10,31] for review), we here study the precision of

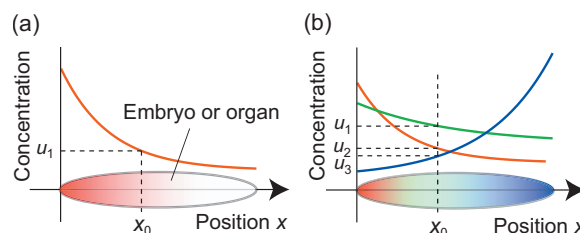


FIG. 1. (Color) Positional specification in one-dimensional space by (a) a single morphogen and (b) by multiple morphogens. Horizontal axis is position x . Each cell senses morphogen concentration(s) and recognizes its own position.

positional information under the situation where cells make best use of multiple noisy inputs in estimating their true positions. To do so, we consider the situation where cells do the maximum-likelihood estimation of the positions based on observed morphogen concentrations.

The formula indicates how the precision can be improved as the number of morphogen species increases. We can estimate the minimum requisite number to achieve a given precision of positional information. Interestingly, the precision also depends on the correlation of noises among different chemical species. We show that if noises added to two chemicals are positively (negatively) correlated, the oppositely (identically) directed gradients give more precise positional information.

We illustrate the use of our formula by applying it to experimental data of spatial distributions of Bicoid and Caudal proteins along the anterior-posterior axis of early development of the *Drosophila* embryo. We estimated the potential accuracy of the positional information given by these two chemicals at each location of an embryo. We found that the combined information provided by the two chemicals is able to give accurate positional information in the middle part of the embryo, where the embryo segmentation occurs in later stages, much more than near both ends.

II. MODEL AND METHOD

A. Basic ideas

Consider the situation in which N diffusive chemicals (morphogens) ($N \geq 1$) are distributed over a one-dimensional space. Denote a vector of their concentrations by $\mathbf{u}(x) \equiv (u_1(x), \dots, u_N(x))$ where $u_i(x)$ is the concentration of morphogen i at location x and is a continuous and smooth function of x . For simplicity, we here assume that for each i ($i = 1, 2, \dots, N$) u_i is a monotonous function of x within a focal spatial region I , which is an interval: $x_A < x < x_P$. The focal region I in the real space is mapped to a curve C in the N -dimensional concentration space (Fig. 2). In the absence of noises in the morphogen concentrations, each spatial point on I has a one-to-one correspondence with a point on this curve C . Hence, in the absence of noises, each cell can correctly recognize its own spatial position from a given \mathbf{u} and can make appropriate responses based on the value. The location estimated from the morphogen concentrations $x(\mathbf{u})$ can be regarded as a mathematical definition of positional information [32].

In contrast, in the presence of noises (any noises including cell-to-cell and embryo-to-embryo variabilities and diffusion noises), the concentrations of morphogens sensed by each cell \mathbf{u}' are random variables and their values may be deviated from the curve C in the concentration space. The estimate of the position \hat{x} from the observed \mathbf{u}' includes errors. In order to achieve normal development, it is desirable for cells to respond in a manner to minimize errors between their true positions and where they think they are. In the case where the cells have no reliable prior information on their positions, the minimization is realized by the maximum-likelihood estimation of the position based on the observed concentrations \mathbf{u}' . Thus we define the positional information

$\hat{x}(\mathbf{u}')$ by the maximum-likelihood estimation in considering the limit of the precision of positional information.

In this paper, we discuss the magnitude of error caused by the maximum-likelihood estimate of the position and how to reduce the error in the estimate under unavoidable noises. Since the main purpose of this study is to examine the amount of information on position given by morphogen gradients, molecular mechanisms in the readout processes are not considered in the following analysis.

B. Mathematical formulation

We denote the morphogen concentrations observed by a cell located at x by \mathbf{u}' . We assume that it follows a multivariate normal distribution with mean equal to $\mathbf{u}(x)$:

$$\Pr[\mathbf{u}'|x] = \frac{1}{\sqrt{(2\pi)^N |\Sigma|}} \exp \left\{ -\frac{1}{2} [\mathbf{u}' - \mathbf{u}(x)]^T \Sigma^{-1} [\mathbf{u}' - \mathbf{u}(x)] \right\}, \quad (1a)$$

where Σ is the variance-covariance matrix of concentration fluctuations around $\mathbf{u}(x)$. It is given as

$$[\Sigma]_{ij} \equiv \begin{cases} \sigma_i^2(u_i), & (i = j) \\ \rho_{ij}(u_i, u_j) \sigma_i(u_i) \sigma_j(u_j), & (i \neq j), \end{cases} \quad (1b)$$

where $\sigma_i^2(u_i)$ is the variance of chemical i , and $\rho_{ij}(u_i, u_j)$ is the correlation coefficient between i and j morphogens.

The maximum-likelihood estimate of the location x (denoted by \hat{x}) obtained from an observed set of the concentrations \mathbf{u}' is the one that maximizes Eq. (1a), and it satisfies $d(\Pr[\mathbf{u}'|x])/dx = 0$. Thus, when cells observe \mathbf{u}' , they are assumed to behave as if they are at $\hat{x}(\mathbf{u}')$. The relation between \mathbf{u}' and \hat{x} defines a map from the concentration space to the real space, by which we can relate the randomness of chemical concentrations sensed by a cell Σ to the precision of the positional information given to the cell $1/\text{Var}[\hat{x}]$. If the variances of concentration fluctuations are small (compared with the curvature radius of C), we can derive a relation between \mathbf{u}' and \hat{x} explicitly. Let us focus on the cell located at x_0 . We assume that the cell does not have additional information other than the morphogen concentrations. We can approximate the curve C by a straight line L around $\mathbf{u}(x_0)$ based on the expansion $\mathbf{u}(x) = \mathbf{u}(x_0) + (x - x_0)(d\mathbf{u}/dx|_0) + \dots$ (see Fig. 2). Then, function $\Pr[\mathbf{u}'|x]$ is maximized when x is equal to

$$\hat{x} = x_0 + \frac{(d\mathbf{u}/dx|_0)^T \Sigma^{-1} [\mathbf{u}' - \mathbf{u}(x_0)]}{(d\mathbf{u}/dx|_0)^T \Sigma^{-1} (d\mathbf{u}/dx|_0)}, \quad (2)$$

where $d\mathbf{u}/dx|_0$ is the gradient vector of morphogen concentrations at x_0 [see Appendix A for the derivation of Eq. (2)]. We also have $E[\hat{x}] = x_0$, implying that the estimate \hat{x} by the cell at x_0 is unbiased. From Eq. (2), all values of \mathbf{u}' satisfying $(d\mathbf{u}/dx|_0)^T \Sigma^{-1} [\mathbf{u}' - \mathbf{u}(x_0)] = \text{const}$ return the same value of \hat{x} . If cells choose their responses (e.g., decision of their fates) based on their estimate \hat{x} then different cells receiving \mathbf{u}' on a plane $(d\mathbf{u}/dx|_0)^T \Sigma^{-1} [\mathbf{u}' - \mathbf{u}(x_0)] = \text{const}$ are expected to behave in a similar manner. This prediction can be used as a test of whether cells might use the maximum-likelihood

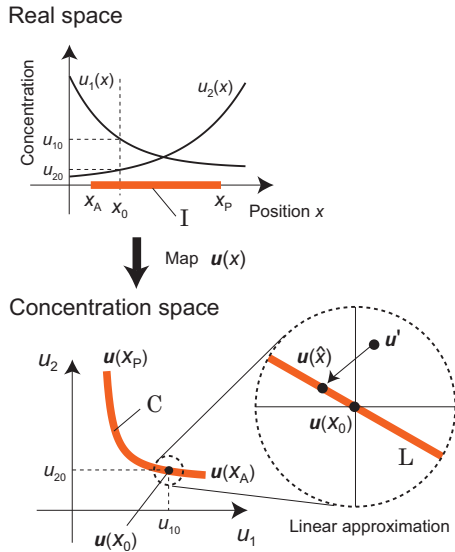


FIG. 2. (Color) Relation between the real space and the concentration space (case of two morphogens). A cell located at x receives a set of morphogen concentrations $\mathbf{u}(x)$. A segment I in the real space is mapped to a curve C in the concentration space by the map $\mathbf{u}(x)$. The dotted circle shows a magnified view around a point $\mathbf{u}(x_0)$ in the concentration space. Around each point, the curve C can be approximated as a line L . In the circle, \mathbf{u}' is a set of concentrations observed by the cell located at x_0 and $\mathbf{u}(\hat{x})$ is the point on L corresponding to the estimated position \hat{x} from \mathbf{u}' . See the text for details.

estimation based on the observed morphogen concentrations.

In this study, we focus on the variance of the positional estimate $\text{Var}[\hat{x}]$, which indicates the magnitude of ambiguity of the positional information provided by noisy morphogen gradients. Its inverse $1/\text{Var}[\hat{x}]$ gives the precision of positional information. According to Appendix B, it is written as follows:

$$\frac{1}{\text{Var}[\hat{x}]} = (\mathbf{d}\mathbf{u}/\mathbf{d}x|_0)^T \Sigma^{-1} (\mathbf{d}\mathbf{u}/\mathbf{d}x|_0). \quad (3)$$

The precision of positional information is determined by the vector of chemical gradients (i.e., $\mathbf{d}\mathbf{u}/\mathbf{d}x$) and the variance-covariance matrix of noises in their concentrations (i.e., Σ). Maximizing the precision corresponds to minimizing the variance of \hat{x} . We have previously discussed the ambiguity of positional information in the case that the spatial dimension is equivalent to the number of chemicals and measured the ambiguity in terms of information entropy $E[\log p(\hat{x})]$, where $p(\hat{x})$ is the probability density function of the positional estimate \hat{x} [32]. In this setting, the entropy in one-dimensional positioning increases with $\log(\text{Var}[\hat{x}])$ when \hat{x} follows a normal distribution. Hence, minimizing the entropy in the previous study [32] is equivalent to minimizing the variance of \hat{x} in the current paper.

C. Positioning performance of a morphogen

In the case of a single morphogen with gradient ($N=1$) [see Fig. 3(a)], the precision of positional information

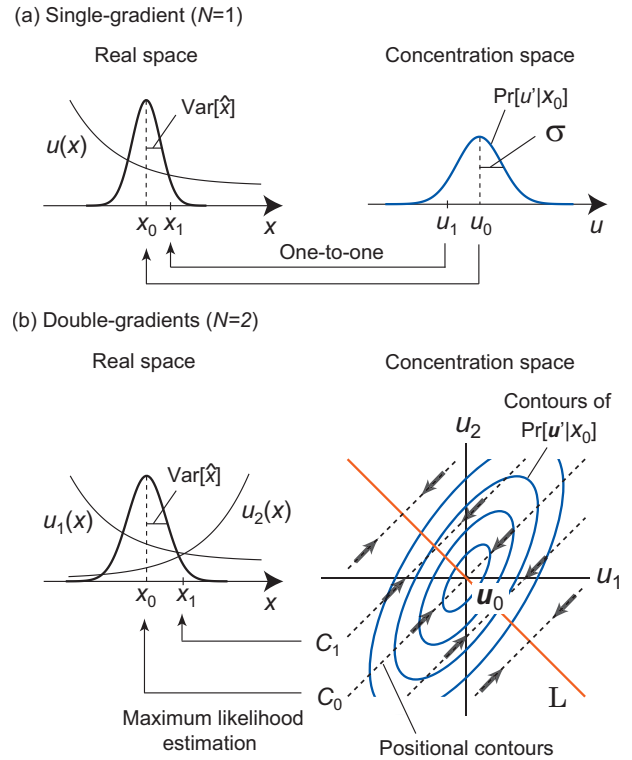


FIG. 3. (Color) Relation between the variance of positional estimate $\text{Var}[\hat{x}]$ and fluctuations in morphogen concentrations. (a) Case of $N=1$. Since each point in the concentration space has a one-to-one correspondence with a point in the real space, $\text{Var}[\hat{x}]$ can be approximated to be proportional to σ^2 . (b) Case of $N=2$. Maximum-likelihood estimation defines a map from the concentration space to the real space (i.e., from \mathbf{u}' to \hat{x} [see Eq. (2)]). In the concentration space, each broken line with arrows shows the contours of positional values; that is, for every observed values of \mathbf{u}' on each broken line (e.g., C_0 and C_1), the estimated position \hat{x} is the same (e.g., x_0 for C_0 and x_1 for C_1). The direction of the contours depends on the values of dPP , η_i , and the correlation of noises in morphogen concentrations ρ_{12} [see Eq. (4)].

$1/\text{Var}[\hat{x}]$ is simply equal to the squared ratio of the gradient steepness du_1/dx to the standard deviation of the noise σ_1 , i.e., $1/\text{Var}[\hat{x}] = [(du_1/dx)/\sigma_1]^2$. Thus $\eta \equiv (du_1/dx)/\sigma_1$ is a good indicator of the performance for the morphogen to specify spatial position. We call η as “directional positioning performance (PP) (dPP)” (see Fig. 4) since η can be either positive or negative, and its sign depends on the direction of the gradient. Negative (or positive) sign corresponds to $du_1/dx < 0$ (or $du_1/dx > 0$). We call the absolute value $|\eta|$ as PP. Its squared value $1/\text{Var}[\hat{x}] = \eta^2$ indicates the precision in specifying the spatial position by the morphogen gradient.

PP is defined at each location x and, in general, it may change with x [33]. Hence, we can discuss the spatial pattern of PP, when the precision of positional information is higher in some parts of an embryo than in other parts. Later, we will discuss this idea with experimental data. In contrast, PP becomes constant in a specific situation where the gradient is exponential and if σ_1 is proportional to the mean $u_1(x)$. For instance, when the average profile of the gradient is given by $u_1(x) = c_1 \exp(a_1 x)$ then $\eta = a_1/\alpha_1$, where α_1 is the linear coefficient of the concentration noise, i.e., $\sigma_1(x) = \alpha_1 u_1(x)$.

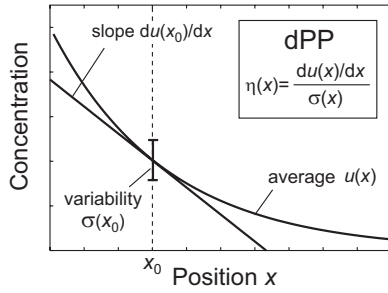


FIG. 4. The definition of dPP of a morphogen. dPP is defined at each location x as the ratio of the gradient $du(x)/dx$ to the standard deviation of the morphogen concentration $\sigma(x)$ at x . The sign of dPP indicates the direction of the gradient.

Note that PP (or its inverse) is a natural measure for the precision (or ambiguity) of positioning given by a single morphogen gradient. The value was calculated for different chemicals in previous studies [14,27,33]. A main purpose of the current study is to extend the measure in the case in which multiple morphogens are operating simultaneously with correlated noises.

III. RESULTS

In this section, we first focus on the case with $N=2$ and derive a formula telling how the precision of positional information depends on the number of morphogens and on the correlation of noises among them. After that, we discuss the case with general N .

A. Case of two morphogens ($N=2$)

In the case of $N=2$, the estimated position \hat{x} for a given observed set of morphogen concentrations \mathbf{u}' is given by

$$\hat{x} = x_0 + \sum_{i=1}^2 w_i \frac{u'_i - u_i(x_0)}{du_i/dx}, \quad (4a)$$

$$w_i = \frac{\eta_i^2 - \rho_{12}\eta_1\eta_2}{\eta_1^2 + \eta_2^2 - 2\rho_{12}\eta_1\eta_2}, \quad (i=1,2), \quad (4b)$$

where $\eta_i \equiv (du_i/dx)/\sigma_i$ is the dPP for morphogen i . Equation (4a) indicates that the estimate \hat{x} is a linear function of observed values u'_i with weighting factor $w_i/(du_i/dx)$, where w_i is determined by the PP values and the correlation of noises between the two morphogens. Equation (4a) defines a map from the concentration space to the real space (i.e., from \mathbf{u}' to \hat{x}). The correspondence is not one to one [cf. the case with $N=1$ (Fig. 3)] because all the points on a line $w_1(du_2/dx)[u'_1 - u_1(x_0)] + w_2(du_1/dx)[u'_2 - u_2(x_0)] = \text{const}$ in the concentration space are mapped to the same point in the real space, suggesting that each of these line defines a positional contours on the concentration space [the broken lines in Fig. 3(b)]. From Eq. (4), the precision of positional estimate $1/\text{Var}[\hat{x}]$ becomes as follows:

$$\frac{1}{\text{Var}[\hat{x}]} = \frac{\eta_1^2 + \eta_2^2 - 2\rho_{12}\eta_1\eta_2}{1 - \rho_{12}^2} \geq \eta_i^2, \quad (i=1,2). \quad (5)$$

The equality holds when $\eta_1 = \rho_{12}\eta_2$ or $\eta_2 = \rho_{12}\eta_1$. The inequality (5) indicates that the precision of positional information provided by two chemicals is greater than that by either one of the two chemicals.

1. Independent noises

Especially, when there is no correlation between noises of the two morphogen concentrations u_1 and u_2 , i.e., $\rho_{12}=0$, we have

$$\frac{1}{\text{Var}[\hat{x}]} = \eta_1^2 + \eta_2^2. \quad (6)$$

Thus, the precision of positional information given by two independent morphogens is the sum of the precision when each morphogen is used ($1/\text{Var}[\hat{x}] = \eta_i^2$ for $N=1$). $1/\text{Var}[\hat{x}]$ may, in general, be a function of spatial position x . However, it is constant if the concentrations of the two morphogens follow exponential distribution and if the standard deviation is proportional to the mean ($\sigma_i \propto u_i$).

2. Correlated noises

Let us come back to the case with correlated noises ($\rho_{12} \neq 0$). Here we show that the precision of positional information can be improved when the noises added to concentrations of two morphogens are correlated. As shown below, the precision $1/\text{Var}[\hat{x}]$ given by Eq. (5) is asymmetric with respect to the sign of the correlation. Depending on the relative direction of two gradients [Figs. 5(a) and 5(b)], the precision is improved more effectively by a positive correlation in one situation but by a negative correlation in the other situation.

Note that the correlation in this paper indicates the one concerning the noises added to concentrations of two morphogens at each location. It is independent of the directions of gradients. Oppositely directed gradients may have a positive correlation of noises, and identically directed gradients may have a negative correlation of noises.

We first consider a case in which the two morphogens have the same PP: $|\eta_1| = |\eta_2|$ [Fig. 5(c) and 1]. In this case, from Eq. (3), the precision is a monotonic function of the correlation coefficient ρ_{12} ; $1/\text{Var}[\hat{x}] = 2\eta^2/(1 - \rho_{12})$ when the gradients are oppositely directed (i.e., $\eta_1 = -\eta_2$) and is $2\eta^2/(1 + \rho_{12})$ when identically oriented (i.e., $\eta_1 = \eta_2$).

Hence, if the correlation is positive ($\rho_{12} > 0$), the two morphogens with the oppositely directed gradients give more accurate positional information than those with the identically directed gradients. In contrast, if the correlation is negative ($\rho_{12} < 0$), two morphogens with identically directed gradients give more accurate positional information.

This result can be explained intuitively in the following manner. Figure 5(d) shows the probability distributions of morphogen concentrations received by two cells located close to each other. When the noises are positively correlated [Figs. 5(d), 1, 5(d), and 2], the overlapped area of the two distributions is smaller when the gradients are oppositely directed than when they are identically directed. If the overlap-

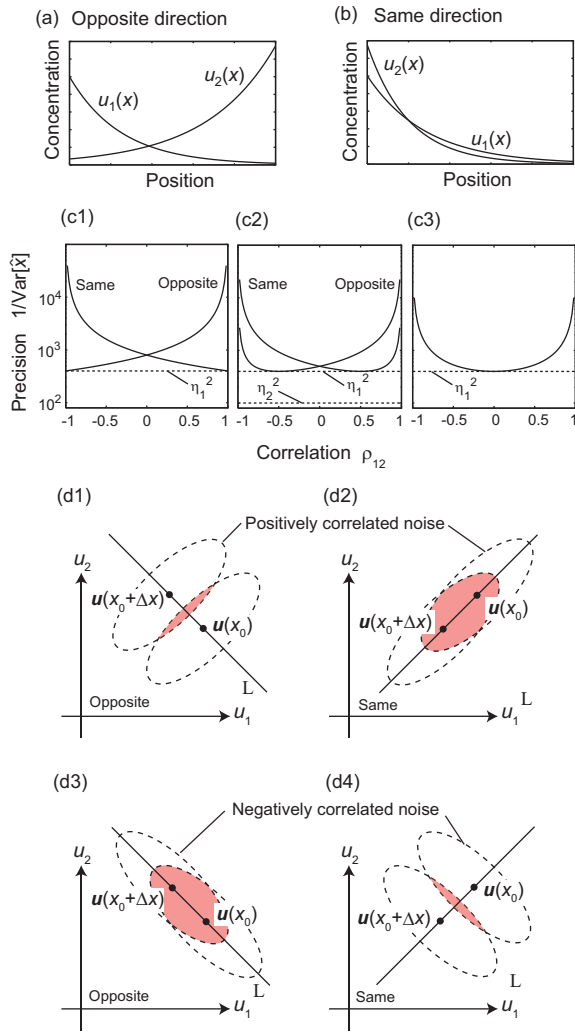


FIG. 5. (Color) The precision of positional information depends on the combination of gradient directions for a given sign of correlation. (a) Oppositely directed gradients. (b) Identically directed gradients. (c) Relation between the correlation of noises ρ_{12} and precision $1/\text{Var}[\hat{x}]$ (semilog plot). (c1) When two chemicals have the same magnitude of positioning performance, the precision becomes a monotonously increasing (decreasing) function of ρ_{12} for the oppositely (identically) directed gradients. Broken line is the precision provided by either of two chemicals. Parameter values: $|\eta_1|=|\eta_2|=20$. (c2) The case where $|\eta_1|\neq|\eta_2|$. Parameter values: $|\eta_1|=20$ and $|\eta_2|=10$. Broken lines are the precision provided by each chemical. (c3) The case where the second chemical has no gradient. Parameter values: $|\eta_1|=20$ and $|\eta_2|=0$. Broken line is the precision provided by the first chemical. (d) Illustration of the reason why the correlation of noises improves the precision of positional information. The ellipsoids indicate the distributions of values of morphogen concentrations received by two cells located at x_0 and $x_0+\Delta x$, respectively. Each value of \mathbf{u}' in the overlapped area of two distributions (shaded in red) is observed by both cells with high probability, implying that the two cells are difficult to distinguish based on the value \mathbf{u}' . If noises are [(d1) and (d2)] positively correlated, the oppositely directed gradients of u_1 and u_2 (note one cell has higher u_1 and lower u_2 than the other cell) lead to the smaller overlapped area and improve the precision of positional estimation. In contrast, in the case of the [(d3) and (d4)] negatively correlated noises, the identically directed gradients can improve the precision.

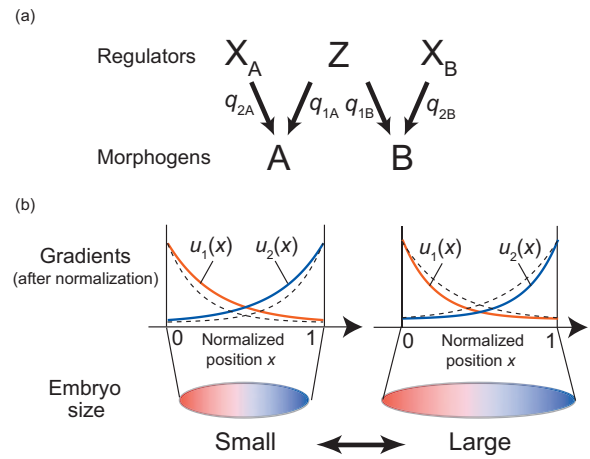


FIG. 6. (Color) Possible mechanisms to generate correlated noises. (a) When the synthesis and/or the degradation of morphogens A and B are regulated by a common factor Z , their concentrations can be correlated. (b) Spatial normalization causes correlation between morphogen concentrations measured at the same relative position. For larger embryos, the concentrations become smaller at the same relative position. See the text for details.

ping is more extensive, it is more difficult to distinguish the two cells based on the observed value \mathbf{u}' . Hence, in the case of positively correlated noises, the oppositely directed gradients leading smaller overlapping can improve the precision of positional estimation from the observed morphogen concentrations. In the extreme case of $\rho \rightarrow 1$, the precision $1/\text{Var}[\hat{x}]$ becomes infinity when the gradients are oppositely directed. From a similar argument, we can conclude that in the case of negatively correlated noises [Figs. 5(d), 3, 5(d), and 4], the identically directed gradients achieve a higher precision than the oppositely directed gradients.

Next, we consider the case in which one signal is more reliable than the other, say $|\eta_1| > |\eta_2|$ [Fig. 5(c) and 2]. In this case, the precision is no longer a monotonic function of the correlation ρ_{12} , but the desirable direction of the gradients depends on the sign of the correlation. If the correlation is positive, oppositely directed gradients are better than the identically directed gradients. But the reverse is true if the correlation is negative.

Interestingly, the correlation can improve the precision even if the value of PP for either chemical $|\eta_1|$ or $|\eta_2|$ is 0, i.e., one of the two morphogens has no gradient [see Fig. 5(c) and 3]. In the presence of the correlation between two noises, the noise of one morphogen without gradient tells the noise of the other with gradient; the latter allowing cells to estimate the position more accurately.

3. Possible mechanisms for correlated noises

The last section shows that the correlation of noises can improve the precision of positional information. Here, we consider possible mechanisms for correlated noises. One scenario is common regulators of synthesis and/or degradation of morphogens [Fig. 6(a)]. As a simple example, let us consider the situation in which the level of morphogen A depends on factor Z and X_A and that of B depends on Z and X_B

as follows: $A=f(Z,X_A)$ and $B=g(Z,X_B)$. Suppose that the amounts of three regulators Z , X_A , and X_B fluctuate around their averages and that they are independent each other. Then, by setting $q_{1A}=\partial f/\partial Z$, $q_{2A}=\partial f/\partial X_A$, $q_{1B}=\partial g/\partial Z$, and $q_{2B}=\partial g/\partial X_B$, we have the following result:

$$\rho_{AB} = \frac{q_{1A}q_{1B}V_Z}{\sqrt{q_{1A}^2V_Z + q_{2A}^2V_{X_A}}\sqrt{q_{1B}^2V_Z + q_{2B}^2V_{X_B}}}, \quad (7)$$

where V_Z , V_{X_A} , and V_{X_B} are the variances of the regulators Z , X_A , and X_B , respectively. Hence, the correlation ρ_{AB} is positive if $q_{1A}q_{1B} > 0$ and is negative if $q_{1A}q_{1B} < 0$. This argument can be extended to the case in which the number of morphogens N is larger than 2.

Second possible mechanism generating the positive correlation in noises is the embryo-to-embryo variability of body size [Fig. 6(b)]. Suppose that to achieve normal development, the position of morphological structures relative to the size of the whole embryo is more important than the absolute distance from an end. We here consider the correlation of fluctuations of two morphogen gradients in a coordinate of relative position. For example, let us consider two exponential gradients given as follows: $u_1(y)=c_1 \exp[-a_1y]$ and $u_2(y)=c_2 \exp[a_2(y-L)]$, where y is the absolute position and L is the size of the focal embryo. c_i and a_i ($i=1,2$) are constants. The location of a cell relative to the whole embryo size is $x=y/L$. Using this, the two gradients are expressed as $u_1(x)=c_1 \exp[-a_1Lx]$ and $u_2(x)=c_2 \exp[a_2L(x-1)]$. If we compare two embryos with different size L , the concentrations of morphogens $u_1(x)$ and $u_2(x)$ at the same relative location x are both smaller for the larger embryo because the decrease rate in gradients are a_1L and a_2L . If the variation in embryo size is much larger than other sources of variations (such as the variance in the source intensity, diffusion coefficient, or decomposition rate), there is a positive correlation between the fluctuations of the two morphogens, irrespective of their directions of gradients.

This correlation should not be interpreted just as an artifact due to the normalization. Rather, in the case where the relative scale is important, the (apparent) correlation can be used to improve the precision of positional information in a coordinate system of relative position. Based on Eq. (5), oppositely directed gradients would effectively reduce the embryo-to-embryo variability of body size. Note that in most studies on positional information, the data of different embryos are compared after the spatial normalization of their sizes by an appropriate typical length (see, for example, [13,14,34]).

B. General case with N morphogens

We can derive similar equations for N morphogens ($N > 2$). When the noises of different morphogen species are independent, the maximum-likelihood estimate of the position \hat{x} for the observed \mathbf{u}' is derived as follows:

$$\hat{x} = x_0 + \sum_{i=1}^N w_i \frac{u'_i - u_i(x_0)}{du_i/dx}, \quad (8)$$

where $w_i = \eta_i^2 / \sum_{i=1}^N \eta_i^2$ is the weight that determines how chemical i contributes to the positional estimation. Then the precision of \hat{x} , $1/\text{Var}[\hat{x}]$, is

$$\frac{1}{\text{Var}[\hat{x}]} = \sum_{i=1}^N \eta_i^2. \quad (9)$$

It is the sum of the squared values of the PP for different morphogens. When all morphogens have the same PP, the precision of positional information is improved in proportion to the number of morphogens N .

The precision of positional information can also be improved in general cases if noises added to chemicals are correlated in an appropriate manner. Here, we consider a case where all correlation coefficients ρ_{ij} ($i \neq j$) are of the same value ρ and all morphogens has the same positioning performance, i.e., $\eta_i = \eta$ for all i . It is noted that ρ has to satisfy $-1/(N-1) < \rho < 1$ when the variance-covariance matrix is positive definite. Then the precision of positional information $1/\text{Var}[\hat{x}]$ becomes as follows (see Appendix C):

$$\frac{1}{\text{Var}[\hat{x}]} = N \left(\frac{1}{1 + (N-1)\rho} \right) \eta^2 \equiv N_{\text{eff}} \eta^2. \quad (10)$$

Equation (10) means that the correlation among morphogens works as a factor to change the effective number of independent morphogens (i.e., from N to N_{eff}). When ρ is negative, the effective number N_{eff} is larger than the real number of morphogens N . In contrast, when ρ is positive, N_{eff} is smaller than N .

C. Biological application

To illustrate how to apply the method proposed in this study, we analyzed the data of the spatial gradients of Bicoid and Caudal proteins along the anteroposterior axis observed in the early development of *Drosophila* embryo. We used the FLYEX database for *Drosophila* embryos [35] because it is one of the few available databases for multiple gradients. Since the data in FLYEX were measured by immunostaining, quantitative comparison between embryos with high precision may be difficult and results may, in general, depend on the details of normalization procedure, as pointed out by Gregor *et al.* [14]. To know the effect of the choice of normalization procedure, we applied the same analyses to three data sets obtained by different normalization procedures (see Appendix D and [36] for details). We confirmed that the results were very similar between different choices of the normalization procedure.

We started by comparing the spatial profile of Bicoid gradient calculated by using FLYEX database with that by Gregor *et al.* who measured the protein level in the absolute (not relative) concentration [see Fig. 7(a)]. The average profile was very similar between the result from FLYEX data and that from Gregor *et al.*'s method. In contrast, the variations are larger in FLYEX data; the standard deviation at each position was 1.5~2.0 times as large as Gregor *et al.*'s data [detailed values depended on the normalization procedures (data not shown)]. Figure 7(b) is the average profile of Caudal calculated by using FLYEX data and error bars indicate standard deviations at each position.

Figure 7(c) shows the spatial profiles of dPP for Bicoid and Caudal. The values of dPP were almost constant in the spatial range where the gradients are steep. Since Bicoid has

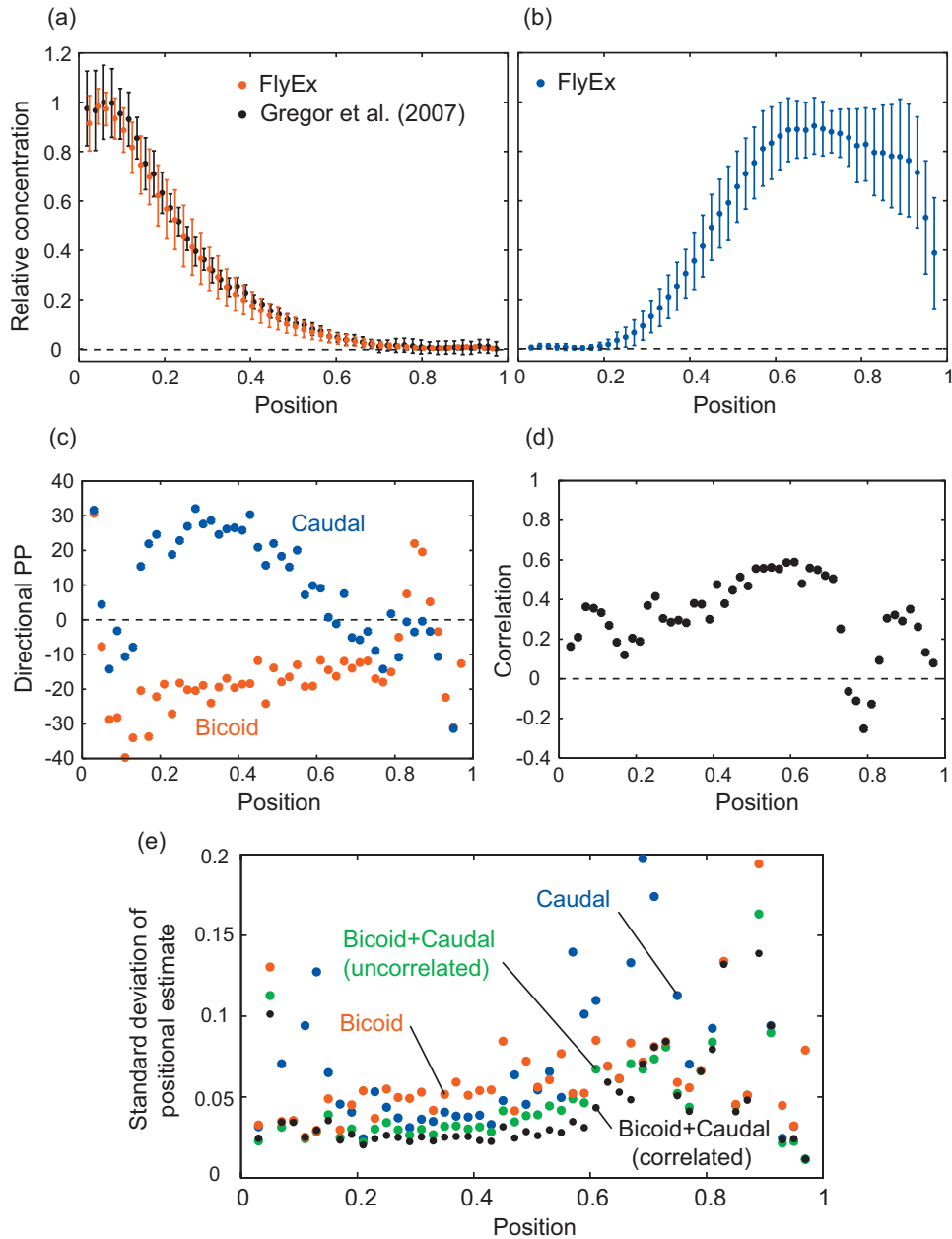


FIG. 7. (Color) Biological application of our formula. By using the FLYEX database on early development of *Drosophila* embryo [35], we calculated spatial profiles of directional positioning performance, correlation, and the precision of positional information for proteins Bicoid and Caudal at the stage 14A1. To know the dependence of results on data processing method, we performed the same analyses for three data sets obtained by different normalization procedures (see Appendix D). Red and blue points are the averaged profile of Bicoid and Caudal over 20 embryos calculated by using a normalization procedure (Appendix D). The red and blue bars show standard deviations. For the purpose of reference, the profile of Bicoid by Gregor *et al.* [14] is shown in (a) (the black points and bars). Average profile was very similar, but, the variations in profile (for Bicoid) calculated for the data from FLYEX are larger than those in Gregor *et al.* (c) Spatial profile of dPP for Bicoid and Caudal. The inverse of squared dPP indicates the precision of positional information provided by each gradient. The sign of dPP indicates the direction of gradient. (d) Spatial profile of the correlation coefficient of noises in the concentrations of Bicoid and Caudal. They are positively correlated in wide area. (e) Standard deviation of positional estimate $\sqrt{\text{Var}[\hat{x}]}$. Red (or blue) dots indicate the standard deviation of positional estimate by using Bicoid (or Caudal) alone. Black (or green) dots are for the case with both Bicoid and Caudal when they are correlated (or independent).

its source at the anterior end and we take the x axis in the posterior direction, the dPP for Bicoid is negative. In contrast, Caudal has its source at the posterior end and Caudal's dPP is positive. As explained before, the squared value of dPP for each chemical determines the precision of positional

information provided by its concentration. In Fig. 7(e), red (or blue) marks indicate the standard deviation of positional estimate, i.e., $\sqrt{\text{Var}[\hat{x}]}$, when only Bicoid (or Caudal) is used. Note that we used the standard deviation to evaluate the precision instead of $1/\text{Var}[\hat{x}]$ because the standard deviation has

a clear biological meaning; that is, it corresponds to the magnitude of ambiguity on the location relative to the whole embryo size. $\sqrt{\text{Var}[\hat{x}]}$ is small in the spatial range where PPs are large. The needs of precise positional specification suggest that the spatial range with high PP value is the region where the chemical works effectively to specify the position.

Next, we calculated the correlation of noises for Bicoid and Caudal at each position. As shown in Fig. 7(d), they are correlated positively in wide area for all three ways of normalization procedures. Due to limited data, we are not able to specify the origins of the correlation clearly. However, we suspect that a large fraction of the correlation is likely to originate from the normalization of the embryonic size as explained in Sec. III A 3. If so, the correlation may be used in an adaptive manner to achieve robust positioning if accuracy of the relative position of the focal point is important.

Black dots in Fig. 7(e) show the spatial profile of the standard deviation of positional information ($\sqrt{\text{Var}[\hat{x}]}$) provided by both Bicoid and Caudal with the correlation calculated above. The precision of positional information is higher ($\sqrt{\text{Var}[\hat{x}]}$ is smaller) in all spatial range than the case if only a single chemical is available (red or blue dots). The precision is also higher in almost all area than the value when the two chemicals are assumed to be independent (green dots). In addition, the precision of positional information changes with the location in an embryo. The precision is higher in the middle part (10–80 %) of the embryo than near both ends of the embryo. This might be related to the fact that the striped gene expression pattern of segmentation genes (e.g., even skipped) occurs in the middle part of the embryo (to be specific, around 25–85 % along the *A-P* axis). But, for more quantitative discussion, we need more data with higher precision.

It should be noted that what we calculated in the above is the amount of information on the position that is provided to the morphogenetic field (i.e., cells along the *A-P* axis) by Bicoid and Caudal gradients themselves, and it is independent of readout mechanisms. Therefore, whether there exist any genes that are directly controlled by both Bicoid and Caudal is not an issue for our question.

IV. DISCUSSION

In this paper, we have derived a formula for the limitation of the precision of positional information provided by the gradients of multiple morphogens with correlated noises. The positional information is defined by an estimate of position from the observed morphogen concentrations $\hat{x}(\mathbf{u}')$, and the precision is defined by the inverse of its variance $1/\text{Var}[\hat{x}]$. Then the “limitation” has been derived under the condition that cells can do the best estimation (i.e., maximum-likelihood estimation) for their positions from the noisy inputs.

According to the formula, the precision increases with the increase in the number of morphogen species adopted. However, considering costs to synthesize more morphogens, it is not necessarily a sensible strategy in biological situations to use too many morphogens in specifying the position. For example, consider the case where an embryo is partitioned into segments and different organs are formed based on the

location of the segments. The estimated location needs to be sufficiently precise so that organogenesis in the subsequent development is performed properly. However, the precision for each cell to recognize its position with the precision finer than a single cell size is unnecessary.

When the correlation among chemical gradients is negligible, from Eq. (9) we can estimate the required number of chemicals N_{eff} to achieve a given precision of positional information $1/\text{Var}[\hat{x}]_0$ as follows:

$$N_{\text{eff}} \approx \frac{1}{\eta^2 \text{Var}[\hat{x}]_0}, \quad (11)$$

where we assume that the PP for each chemical has the same value $|\eta|$. Chemical species with much smaller PP values hardly contribute to the precision. We can interpret N_{eff} in Eq. (11) as the minimum effective number of independent chemicals needed to achieve the precision given by $1/\text{Var}[\hat{x}]_0$.

In addition to the number of morphogens, the relative direction of morphogen gradients affects the precision of positional information. An appropriate choice of the location of morphogen sources is able to improve the precision of positional information. Specifically, gradients of the opposite directions give more accurate information when the noises are positively correlated, while gradients of the same directions do better when the noises are negatively correlated. This result is an extension of our previous study on the optimal placement of multiple morphogen sources [32], where the ambiguity of positional information was measured by using the information entropy for $p(\hat{x})$ and it was proposed that morphogen sources should be placed so as to minimize the positional ambiguity (entropy) over a given target region.

In this study, we have focused on the positional specification in a one-dimensional space to show our basic idea clearly. However, the mathematical formulation discussed here can be extended to the case of multidimensional positional specification. We assumed that each cell infers its position independent of the behaviors of other cells. It is a difficult but important future theoretical problem how the precision can be improved when the intercellular interaction is included in the estimation of positions.

Further, all results in this paper are calculated in the steady state. We think that the quasiequilibrium assumption is a good approximation, at least, as the first step of research because the time scale of readout of chemical gradients is faster than that of the organ growth and change in source levels. It is also a major challenge to extend our theory to the case in the nonsteady state. For example, the source levels of some diffusive chemicals observed in *Drosophila* development are known to change over time [28,37,38].

Recent advancement of experimental techniques has enabled us to measure the variability of multiple chemical gradients. In order to extract essential features of systems from the huge data, we need indicators quantifying the features well. Directional positioning performance η and the precision of positional estimate $1/\text{Var}[\hat{x}]$ may be good examples of such indicators. Such quantitative analysis will shed light on the design of robust developmental processes.

ACKNOWLEDGMENTS

This work was done by the support of a PRESTO procedure of JST to Y.M. and a Grant-in-Aid of JSPS to Y.I. We thank the following people for their very helpful comments: K. Fujimoto, S. Ishihara, K. Kaneko, and A. Nakajima.

APPENDIX A: DERIVATION OF Eq. (2)

The maximum-likelihood estimate of position \hat{x} for the observed set of chemical concentrations \mathbf{u}' satisfies $\partial(\Pr[\mathbf{u}'|x])/ \partial x = 0$, where $\Pr[\mathbf{u}'|x]$ is the likelihood function described by Eq. (1a). When the fluctuations of chemical concentrations are small, $\Pr[\mathbf{u}'|x]$ can be approximated as follows:

$$\Pr[\mathbf{u}'|x] = \frac{1}{(2\pi)^{N/2} \sqrt{|\Sigma(x)|}} \exp\left[-\frac{1}{2} \mathbf{g}^T \Sigma^{-1} \mathbf{g}\right], \quad (\text{A1})$$

where $|\Sigma(x)|$ is the determinant of the variance-covariance matrix $\Sigma(x)$, and $\mathbf{g} = \mathbf{u}' - \mathbf{u}(x_0) - (x - x_0)(d\mathbf{u}/dx|_0)$. Since the concentration \mathbf{u}' observed by the cell located at x_0 obeys $\Pr[\mathbf{u}'|x_0]$, $E[\mathbf{u}' - \mathbf{u}(x_0)] = \mathbf{0}$ holds. At the optimal value of x , we have $(d/dx)\ln(\Pr[\mathbf{u}'|x]) = 0$, which leads to

$$-\frac{1}{2|\Sigma(x)|} \frac{d|\Sigma(x)|}{dx} - \frac{1}{2} \mathbf{g}^T \frac{d\Sigma^{-1}(x)}{dx} \mathbf{g} + \left(\frac{d\mathbf{u}}{dx}\bigg|_0\right)^T \Sigma^{-1}(x) \mathbf{g} = 0. \quad (\text{A2})$$

When the order of the magnitude of concentration noise is ϵ , i.e., $\sigma_i = O(\epsilon)$, the orders of magnitude of the first and second terms in the left-hand side in Eq. (A2) are 1 and that of the third term is $1/\epsilon$. Thus, we can neglect the first and second terms in Eq. (A2). Furthermore, considering $\Sigma^{-1}(x) \approx \Sigma^{-1}(x_0)$, we have

$$(d\mathbf{u}/dx|_0)^T \Sigma^{-1}(x_0) \mathbf{g} = 0. \quad (\text{A3})$$

Since Eq. (A3) is a linear equation of x , the estimated position satisfying Eq. (A3) \hat{x} is uniquely obtained.

APPENDIX B: DERIVATION OF Eq. (3)

Using Eq. (2), the variance is calculated as

$$\text{Var}[\hat{x}] = E[(\hat{x} - x_0)^2], \quad (\text{B1})$$

$$= E\left[\left(\frac{D_u^T \Sigma^{-1} \delta_u}{D_u^T \Sigma^{-1} D_u}\right)^2\right], \quad (\text{B2})$$

$$= \frac{E[(D_u^T \Sigma^{-1} \delta_u)(\delta_u^T \Sigma^{-1} D_u)]}{(D_u^T \Sigma^{-1} D_u)^2}, \quad (\text{B3})$$

where $D_u = d\mathbf{u}/dx|_0$ and $\delta_u = \mathbf{u}' - \mathbf{u}(x_0)$. Noting $\Sigma = E[\delta_u \delta_u^T]$, the numerator of Eq. (B3) becomes

$$[\text{Numerator}] = D_u^T \Sigma^{-1} \Sigma \Sigma^{-1} D_u, \quad (\text{B4})$$

$$= D_u^T \Sigma^{-1} D_u. \quad (\text{B5})$$

Then Eq. (B3) becomes Eq. (3) in the text.

APPENDIX C: DERIVATION OF Eq. (10)

When all correlation coefficients ρ_{ij} ($i \neq j$) are equal to the same value ρ , the inverse matrix of the variance-covariance matrix Σ becomes

$$\Sigma_N^{-1} = \begin{pmatrix} \frac{C_N}{\sigma_1^2} & \frac{D_N}{\sigma_1 \sigma_2} & \dots & \frac{D_N}{\sigma_1 \sigma_N} \\ \frac{D_N}{\sigma_1 \sigma_2} & \frac{C_N}{\sigma_2^2} & \dots & \frac{D_N}{\sigma_2 \sigma_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{D_N}{\sigma_1 \sigma_N} & \frac{D_N}{\sigma_2 \sigma_N} & \dots & \frac{C_N}{\sigma_N^2} \end{pmatrix}, \quad (\text{C1})$$

$$C_N = \frac{1 + (N-2)\rho}{1 + (N-2)\rho - (N-1)\rho^2}, \quad (\text{C2})$$

$$D_N = \frac{-\rho}{1 + (N-2)\rho - (N-1)\rho^2}. \quad (\text{C3})$$

By using the relation $du_i/dx = \eta_i \sigma_i$, we have

$$\frac{1}{\text{Var}[\hat{x}]} = \sum_{i=1}^N \eta_i^2 + A_N \sum_{i=1}^N \eta_i^2 + B_N \sum_{i < j} \eta_i \eta_j, \quad (\text{C4})$$

$$A_N = \frac{(N-1)\rho^2}{1 + (N-2)\rho - (N-1)\rho^2}, \quad (\text{C5})$$

$$B_N = \frac{-2\rho}{1 + (N-2)\rho - (N-1)\rho^2}. \quad (\text{C6})$$

In Eq. (C4), the first term in the right-hand side is the effect of the number of diffusive chemicals on the precision of positional information, and the second and third terms are the effect of the correlation among them. By substituting $\eta_i = \eta$ for all i , we obtain Eq. (10).

APPENDIX D: DATA ANALYSIS FOR THE GRADIENTS OF BICOID AND CAUDAL PROTEINS IN DROSOPHILA EMBRYO

We used the data of protein levels for Bicoid and Caudal at stage 14A1 (20 embryos whose name begins with the letter ‘‘a.’’ or ‘‘cb.’’) from database FLYEX [35]. The stage 14A1 is well before the segmentation process (including the formation of the striped pattern of the even-skipped gene expression). The protein levels were measured by using fluorescently tagged antibodies. In order to adjust the variability of the fluorescent intensity among embryos, we normalized the data of each embryo. To know the effect of the choice of the normalization procedure on results, we applied the same analysis to three data sets obtained by the following normalization procedures (N1–N3).

(i) (N1) Background noise was subtracted from raw data for each embryo according to the method explained in [36]. This kind of data is called ‘‘without background’’ in the database.

(ii) (N2) Background noise was subtracted as explained in (N1), after which the data for each embryo were divided by the maximum intensity in the embryo.

(iii) (N3) Background noise was subtracted as explained in (N1), after which the data for each embryo were divided by a constant A_n which were obtained as the values minimizing the sum of the square deviations,

$$\chi^2 = \sum_{n=1}^N \int |I_n(x) - A_n \bar{c}(x)|^2 dx, \quad (\text{D1})$$

$I_n(x)$ is the fluorescent intensity at position x and A_n is the normalization constant for embryo n . $\bar{c}(x)$ is the average pro-

file of the chemical. We numerically derived A_n and $\bar{c}(x)$ based on the variation method.

In the statistical analyses, we used the data included in 45–55% stripe (called as 10% stripe in the database) perpendicular to the dorsoventral axis, and the space is divided into 50 bins along the anteroposterior axis. The spatial profiles of the mean and variance of the gradients are calculated for the data included in each bin. The correlation of noises included in Bicoid and Caudal gradients was calculated at each location over the data of all pairs for Bicoid and Caudal within each bin. As stated in the text, we confirmed that the results were very similar irrespective of the choice of the normalization procedures. The data shown in Fig. 7 are those obtained by the procedure (N2).

-
- [1] L. Wolpert, *Principles of Development* (Oxford University Press, New York, 1998).
- [2] A. A. Teleman and S. M. Cohen, *Cell* **103**, 971 (2000).
- [3] A. A. Teleman, M. Strigini, and S. M. Cohen, *Cell* **105**, 559 (2001).
- [4] M. Strigini and S. M. Cohen, *Curr. Biol.* **10**, 293 (2000).
- [5] A. D. Lander, Q. Nie, and F. Y. M. Wan, *Dev. Cell* **2**, 785 (2002).
- [6] T. B. Kornberg and A. Guha, *Curr. Opin. Genet. Dev.* **17**, 264 (2007).
- [7] M. Gonzalez-Gaitan, *Natl. Prod. Rev.* **4**, 213 (2003).
- [8] L. Wolpert, *J. Theor. Biol.* **25**, 1 (1969).
- [9] J. B. Gurdon and P. Y. Bourillot, *Nature (London)* **413**, 797 (2001).
- [10] H. L. Ashe and J. Briscoe, *Development* **133**, 385 (2006).
- [11] M. Freeman, *Nature (London)* **408**, 313 (2000).
- [12] J. Jaeger, D. Irons, and N. Monk, *Development* **135**, 3175 (2008).
- [13] B. Houchmandzadeh, E. Wieschaus, and S. Leibler, *Nature (London)* **415**, 798 (2002).
- [14] T. Gregor, D. W. Tank, E. F. Wieschaus, and W. Bialek, *Cell* **130**, 153 (2007).
- [15] T. Gregor, E. F. Wieschaus, A. P. McGregor, W. Bialek, and D. W. Tank, *Cell* **130**, 141 (2007).
- [16] S. G. Megason and S. E. Fraser, *Cell* **130**, 784 (2007).
- [17] J. Paulsson, *Nature (London)* **427**, 415 (2004).
- [18] T. Shibata and K. Fujimoto, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 331 (2005).
- [19] Y. Morishita, T. J. Kobayashi, and K. Aihara, *J. Theor. Biol.* **235**, 241 (2005).
- [20] Y. Morishita, T. J. Kobayashi, and K. Aihara, *Biophys. J.* **91**, 2072 (2006).
- [21] B. Houchmandzadeh, E. Wieschaus, and S. Leibler, *Phys. Rev. E* **72**, 061920 (2005).
- [22] T. Bollenbach, K. Kruse, P. Pantazis, M. Gonzalez-Gaitan, and F. Julicher, *Phys. Rev. Lett.* **94**, 018103 (2005).
- [23] M. Howard and P. Rein ten Wolde, *Phys. Rev. Lett.* **95**, 208103 (2005).
- [24] T. Aegerter-Wilmsen, C. M. Aegerter, and T. Bisseling, *J. Theor. Biol.* **234**, 13 (2005).
- [25] A. Eldar, D. Rosin, B.-Z. Shilo, and N. Barkai, *Dev. Cell* **5**, 635 (2003).
- [26] R. A. Veitia and H. F. Nijhout, *BioEssays* **28**, 282 (2006).
- [27] F. Tostevin, P. R. ten Wolde, and M. Howard, *PLOS Comput. Biol.* **3**, e78 (2007).
- [28] J. Jaeger and J. Reinitz, *BioEssays* **28**, 1102 (2006).
- [29] S. Gilbert, *Developmental Biology*, 8th ed. (Sinauer Associates, Sunderland, CT, 2006).
- [30] A. Spirov, K. Fahmy, M. Schneider, E. Frei, M. Noll, and S. Baumgartner, *Development* **136**, 605 (2009).
- [31] G. T. Reeves, C. B. Muratov, T. Schupbach, and S. Y. Shvartsman, *Dev. Cell* **11**, 289 (2006).
- [32] Y. Morishita and Y. Iwasa, *Phys. Rev. E* **77**, 041909 (2008).
- [33] T. Bollenbach, P. Pantazis, A. Kicheva, C. Bokel, M. Gonzalez-Gaitan, and F. Julicher, *Development* **135**, 1137 (2008).
- [34] J. Jaeger, S. Surkova, M. Blagov, H. Janssens, D. Kosman, K. N. Kozlov, Manu, E. Myasnikova, C. E. Vanario-Alonso, M. Samsonova, D. H. Sharp, and J. Reinitz, *Nature (London)* **430**, 368 (2004).
- [35] <http://flyex.ams.sunysb.edu/FlyEx/>
- [36] E. Myasnikova, M. Samsonova, D. Kosman, and J. Reinitz, *Dev. Genes Evol.* **215**, 320 (2005).
- [37] N. Yakovy, J. Lembong, T. Schupbach, and S. Y. Shvartsman, *Development* **135**, 343 (2008).
- [38] S. Bergmann, O. Sandler, H. Sberro, S. Shnider, E. Schejter, B.-Z. Shilo, and N. Barkai, *PLOS Comput. Biol.* **5**, e46 (2007).