Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 October 2019 doi:10.20944/preprints201910.0266.v1

Peer-reviewed version available at *Biomimetics* **2020**, *5*, 10; doi:10.3390/biomimetics5010010

Article

Optimal Sensing for Fish School Identification

Pascal Weber [1], Georgios Arampatzis [1],Guido Novati [1], Siddhartha Verma [2], Costas Papadimitriou [3] and Petros Koumoutsakos [1,*]

[1] Computational Science and Engineering Laboratory, Clausiusstrasse 33, ETH Zürich, CH-8092, Switzerland.
[2] Department of Ocean and Mechanical Engineering, Florida Atlantic University, Boca Raton, FL 33431, USA and Harbor Branch Oceanographic Institute, Florida Atlantic University, Fort Pierce, FL 34946, USA
[3] Department of Mechanical Engineering, University of Thessaly, Pedion Areos, GR-38334 Volos, Greece
* Correspondence: petros@ethz.ch

Abstract: Fish schooling implies an awareness of the swimmers for their companions. In flow mediated environments, in addition to visual cues, pressure and shear sensors on the fish body are critical for providing quantitative information that assists the quantification of proximity to other swimmers. Here we examine the distribution of sensors on the surface of an artificial swimmer so that it can optimally identify a leading group of swimmers. We employ Bayesian experimental design coupled with two-dimensional Navier Stokes equations for multiple self-propelled swimmers. The follower tracks the school using information from its own surface pressure and shear stress. We demonstrate that the optimal sensor distribution of the follower is qualitatively similar to the distribution of neuromasts on fish. Our results show that it is possible to identify accurately the center of mass and even the number of the leading swimmers using surface only information.

Keywords: Bayesian experimental design; optimal sensor placement; schooling; self-propelled swimmers; lateral line

1. Introduction

Fish navigate in their environments by processing and adapting to cues from the surrounding flow fields. The flow environment is replete with mechanical disturbances (pressure, vorticity) that can convey information about the sources that generated them. Fish swimming in groups have been found to process such external information and balance it with social cues [1,2]. Fish may perceive mechanical disturbances in terms of surface pressure and shear stresses through neuromasts. Early studies and experiments showed that the functioning of the lateral line is crucial for several tasks [3,4]. Experiments with trouts in the vicinity of objects showed its importance for Kármán gaiting and bow wake swimming as well as energy efficient station keeping [5,6] Additionally, it was shown that the flow information allows to determine the cylinder diameter and flow velocity as well as the position relative to the generated Kármán vortex street by taking measurements of the flow [7,8]. Using blind cave fish a study showed the importance of the lateral line to detect location and shape of surrounding objects and avoid obstacles [9–12]. In another study the feeding behavior of blinded mottled sculpin was tested and it was found that they use their lateral line system to detect prey [13]. Also it was found that blind fish manage to keep their position in schools and lose this ability with disabled lateral line organ [14]. Further importance of the lateral line was shown for communicational behavior [15], selection of habitats [16] and rheotaxis [17].

The lateral line consists of two functional classes of receptors: mechanoreceptors and electroreceptors. In this work we mimic the mechanosensory receptors, more specifically the sub-surface 'canal' neuromasts and superficial neuromasts [18,19]. The neuromast on the fish skin are used to detect shear stresses, where the ones residing in the lateral line canals are used to detect pressure

© 2019 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 October 2019 doi:10.20944/preprints201910.0266.v1

Peer-reviewed version available at Biomimetics 2020, 5, 10; doi:10.3390/biomimetics5010010

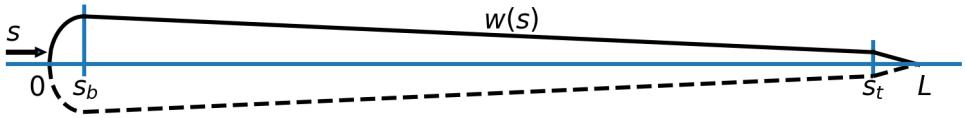


Figure 1. Parametrization of the fish surface as described in Eq. (1).

gradients [20–24]. Due to the filtering nature of the canals, the detection of small hydrodynamic stimuli against background noise is improved for the subsurface neuromasts [25].

The effectiveness and versatility of the lateral line organ inspired several bio-inspired artificial flow sensors [26–29]. Arranging these sensors in arrays on artificial swimmers has attracted attention to transform underwater sensing [30–38]. Here, leveraging the intelligent distributed sensing inspired by the lateral line showed to be effective in robots moving in aquatic environments [39–41].

In order to better use and understand the capabilities of the artificial sensors several studies regarding the information content in the flow and optimal harvesting of this information were performed: The prevalence of information on the position of a vibrating source was shown to be linearly coded in the pressure gradients measured by the subsurface neuromasts [42]. Furthermore, it was shown that the variance of the pressure gradient correlated with the presence of lateral line canals [43]. A study with fish robots equipped with distributed pressure sensors for flow sensing and Bayesian filtering for estimating the flow speed, angle of attack, and foil camber [44]. Other studies focused on dipole sources to develop methods to extract information and optimize the parameters of the sensing devices [45,46]. In a recent study artificial neural networks were employed to classify wakes from airfoils [47]. In order to find effective sensor positions weight analysis algorithms were employed [48].

In this study, following earlier work for detection of flow disturbances by single obstacles, [49] we examine the optimality of the spatial distribution of sensors in a self-propelled swimmer who infers the number in a group of neighboring (here leading) fish and their relative positions. We combine numerical simulations of the two-dimensional Navier-Stokes equation and Bayesian optimal sensor placement to examine the extraction of flow information by pressure gradients and optimal positioning of sensors. The paper is organised as follows: In Section 2.1 we describe the numerical simulations and in Section 2.2 the process of Bayesian optimal experimental design. We present our results in Section 3 and conclude in Section 4.

2. Materials and Methods

2.1. Flow simulations

The swimmers are modeled by slender deforming bodies of length L which are characterized by their half-width $w(s)$ along the midline [50,51]

$$w(s) = \begin{cases} \sqrt{2w_h s - s^2}, & 0 \le s < s_b \\ w_h - (w_h - w_t)\left(\frac{s-s_b}{s_t-s_b}\right), & s_b \le s < s_t \\ w_t \frac{L-s}{L-s_t}, & s_t \le s \le L\,. \end{cases} \tag{1}$$

A sketch of the parametrization is presented in Fig. 1. Following [52], we use $w_h = s_b = 0.04L$, $s_t = 0.95L$ and $w_t = 0.01L$. The swimmers propel themselves by performing sinusoidal undulations of their midline. This motion is described by a prescribed time-dependent parameterization of the curvature,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 October 2019 doi:10.20944/preprints201910.0266.v1

Peer-reviewed version available at Biomimetics 2020, 5, 10; doi:10.3390/biomimetics5010010

$$k(s,t) = A(s)\sin\left(\frac{2\pi t}{T_p} - \frac{2\pi s}{L}\right). \tag{2}$$

Here $T_p = 1$ is the tail-beat period and A is the undulation amplitude which linearly increases from $A(0) = 0.82/L$ to $A(L) = 5.7/L$ to replicate the anguilliform swimming motion described by [53]. Given the curvature along s and a center of mass, the coordinates $\mathbf{r}(s,t)$ of the swimmer's midline can be computed by integrating the Frenet–Serret formulas [52]. In turn, the half-width $w(s)$ and the coordinates $\mathbf{r}(s,t)$ characterize the swimmer's surface.

The flow environment is described by numerical simulations of the two-dimensional incompressible Navier-Stokes equations (NSE) in velocity-pressure ($\mathbf{u} - p$) formulation. The NSE are discretized with second order finite differences and integrated in time with explicit Euler time stepping. The fluid-structure interaction is approximated with Brinkman penalization [54–56] by extending the fluid velocity $\mathbf{u}$ inside the swimmers' bodies and by including in the NSE a penalization term to enforce no-slip and no-through boundary conditions,

$$\frac{\mathbf{u}^{k+1} - \mathbf{u}^k}{\delta t} = -\nabla p^k - (\mathbf{u}^k \cdot \nabla)\mathbf{u}^k + \nu\Delta\mathbf{u}^k + \sum_{i=1}^{N_s} \lambda\chi_i(\mathbf{u}^k_{s,i} - \mathbf{u}^k). \tag{3}$$

Here, ν is the kinematic viscosity, $\lambda = 1/\delta t$ is the penalization coefficient, N_s is the number of swimmers, $\mathbf{u}^k_{s,i}$ is the velocity field imposed by swimmer i (composed of translational, rotational and undulatory motions), and χ_i is its characteristic function which takes value 1 inside the body of swimmer i and value 0 outside. The characteristic function χ_i is computed, given the distance of each grid-point from the surface of swimmer i, by a second-order accurate finite difference approximation of a Heaviside function [57]. The pressure field is computed by pressure-projection [56,58],

$$\Delta p = \frac{1}{\delta t}\nabla\cdot\tilde{\mathbf{u}}^k - \frac{1}{\delta t}\sum_{i=1}^{N_s}\chi_i\nabla\cdot\tilde{\mathbf{u}}^k, \tag{4}$$

where $\tilde{\mathbf{u}}^k = \mathbf{u}^k - (\mathbf{u}^k\cdot\nabla)\mathbf{u}^k + \nu\Delta\mathbf{u}^k$. The terms inside the summation in Eq. (4) are due to the non-divergence free deformation of the swimmers.

2.1.1. Schooling formation

The tail-beating motion that propels forwards a single swimmer generates in its wake a sequence of vortices. The momentum contained in the flow field induces forces to the swimmers that fish in schooling formation must overcome to maintain their positions in the group [59]. In this study, we maintain the schooling formation for multiple swimmers by employing closed-loop parametric controllers. We increase or decrease the tail-beating frequency $T_{p,i}$ of each swimmer i if it lags behind or surpasses a desired position in the direction of the school's motion,

$$T_{p,i} = T_p(1 - \Delta x_i). \tag{5}$$

In order to straighten the school's trajectory we impose additional curvature $k_{C,i}$ along each swimmer's midline in order to minimize its lateral deviation Δy_i and its angular deflection $\Delta\theta_i$,

$$k_{C,i} = [\Delta y_i, \langle\Delta\theta_i\rangle]_- + [\langle\Delta y_i\rangle, \Delta\theta_i]_- + [\langle\Delta y_i\rangle, \langle\Delta\theta_i\rangle]_-. \tag{6}$$

Here, $\langle\cdot\rangle$ defines an exponential moving average with weight $\delta t/T_p$, which approximates the integral term found in PI controllers and,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 October 2019 doi:10.20944/preprints201910.0266.v1

Peer-reviewed version available at Biomimetics 2020, 5, 10; doi:10.3390/biomimetics5010010

$$[a,b]_- = \begin{cases} \|a\|\, b & \text{if } ab < 0, \\ 0 & \text{otherwise}\,. \end{cases} \tag{7}$$

The formulation in Eq. (6) indicates that if both the lateral displacement and the angular deviation are positive (or both negative) the swimmer will gradually revert to its position in the formation. Conversely, if Δy_i and $\Delta\theta_i$ have different signs the displacement has to be corrected by adding (or subtracting) curvature to the swimmer's midline.

2.1.2. Flow sensors

We distinguish two types of sensors on the swimmer body. The superficial neuromasts detect flow stress and the subcanal neuromasts pressure gradients [28,60,61]. From the numerical solution of the 2D Navier-Stokes equation we obtain the flow velocity $\mathbf{u} = (u, v)$ and the pressure p at every point of the computational grid. The surface values of these quantities are obtained through a bi-linear interpolation from the nearest grid points. We perform offline analysis by recording the interpolated pressure p and flow velocity $\mathbf{u}$ in the vicinity of the body. We remark that we have neglected points near the end of the body to reduce the influence of large flow gradients that are generated by the motion and sharp geometry of the tail. The shear stresses are computed on the body surface using the local tangential velocity in the two nearest grid points. Moreover we compute pressure gradients along the surface by first smoothing thes pressure along the surface a using splines implemented in SciPy [62,63].

2.2. Optimal sensor placement based on Information gain

In the present work a swimmer is equipped with sensors that are used to identify the size and location of a nearby school. The optimal sensor locations are identified using Bayesian experimental design [64] so that the information obtained from the collected measurements is maximized. We define the information gain as the distance between the prior belief on the quantities of interest and the posterior belief after obtaining the measurements. Here, we choose as measure of the distance the Kullback-Leibler divergence between the prior and the posterior distribution.

2.2.1. Bayesian Estimation of Swimmers

In the present experiment setup, we consider a group of swimmers followed by a single swimmer. The follower needs to identify (i) the relative location $\mathbf{r}$ of the center of mass and (ii) the population n_f of the leading group. We denote with $\boldsymbol{\vartheta} = \mathbf{r}$ or $\boldsymbol{\vartheta} = n_f$ these unknown quantities and allow the follower to update its prior belief $p(\boldsymbol{\vartheta})$ about the leading group of fish by collecting measurements on its sensors. These sensors are distributed symmetrically on both sides of the fish and can be represented by a single point on its mid-line. We denote the k-th measurements location at the upper and the lower part with $\boldsymbol{x}_1(s_k)$ and $\boldsymbol{x}_2(s_k)$, respectively (see Fig. 2 for a sketch of the setup).

We denote by $\boldsymbol{F}(\boldsymbol{\vartheta}; \boldsymbol{s}) \in R^{2n}$ the flow output and include an error term $\boldsymbol{\varepsilon}$ to account for inaccuracies such as as numerical errors and imperfections in the sensors. The measurements on the fish body can be expressed as,

$$\boldsymbol{y} = \boldsymbol{F}(\boldsymbol{\vartheta}; \boldsymbol{s}) + \boldsymbol{\varepsilon}\,. \tag{8}$$

We model the error term by a multivariate Gaussian distribution $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}(\boldsymbol{s}))$ with zero mean and covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{s}) \in \mathbb{R}^{2n\times 2n}$ so that the likelihood of a measurement is expressed as,

$$p\left(\boldsymbol{y}|\boldsymbol{\vartheta}, \boldsymbol{s}\right) = \frac{1}{\sqrt{(2\pi)^{2n}\det(\boldsymbol{\Sigma}(\boldsymbol{s}))}} \exp\left(-\frac{1}{2}\left(\boldsymbol{y} - \boldsymbol{F}(\boldsymbol{\vartheta}; \boldsymbol{s})\right)^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{s})\left(\boldsymbol{y} - \boldsymbol{F}(\boldsymbol{\vartheta}; \boldsymbol{s})\right)\right)\,. \tag{9}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 October 2019 doi:10.20944/preprints201910.0266.v1

Peer-reviewed version available at Biomimetics 2020, 5, 10; doi:10.3390/biomimetics5010010

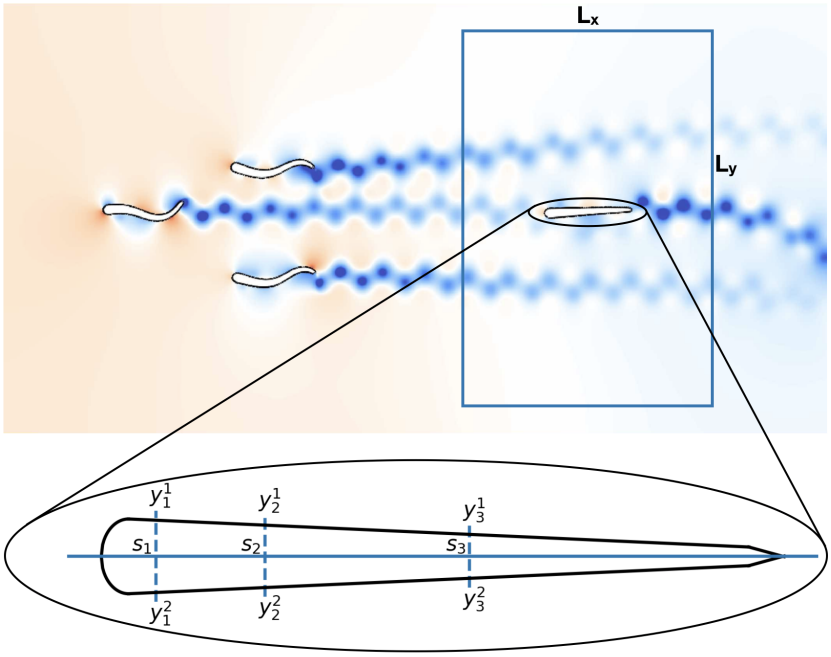


Figure 2. Simulation setup used for determining the optimal sensor distribution on a fish-like body. The follower is initially located inside the rectangular area. The number of fish in the leading group is varied between one and eight. The sensor-placement algorithm attempts to find the arrangement of sensors $\boldsymbol{s}$ that allows the follower to determine with lowest uncertainty the relative position $\mathbf{r}$ and the number of fish n_f in the leading group of fish. For each sensor s_i the fish collects measurements y_i^1 and y_i^2 at locations $\boldsymbol{x}_1(s_i)$ and $\boldsymbol{x}_2(s_i)$ on the skin, respectively.

The covariance matrix depends on the sensor positions $\boldsymbol{s}$ and we assume that the prediction errors are correlated for measurements on the same side of the swimmer and uncorrelated if they originate from opposite sides. Finally, we assume that the correlation is decaying exponentially with the distance of the measurement locations. The functional form of the resulting covariance matrix is given by,

$$\Sigma_{ij}(\boldsymbol{s}) = \begin{cases} \sigma^2 \exp\left(-\frac{\|\boldsymbol{x}_1(s_i)-\boldsymbol{x}_1(s_j)\|}{\ell}\right), & \text{if } 1 \leq i,j \leq n\,, \\ \sigma^2 \exp\left(-\frac{\|\boldsymbol{x}_2(s_{i-n})-\boldsymbol{x}_2(s_{j-n})\|}{\ell}\right), & \text{if } n < i,j \leq 2n\,, \\ 0 & \text{otherwise}\,. \end{cases} \tag{10}$$

where $\ell > 0$ is correlation length and σ is the correlation strength. For all the cases described in this work, the correlation length is set to one tenth of the fish length $\ell = 0.1L$. The correlation strength is set to be two times the mean over the signal coming from our simulations:

$$\sigma = \frac{1}{n\,N_{\boldsymbol{\vartheta}}} \sum_{j=1}^{2n} \sum_{i=1}^{N_{\boldsymbol{\vartheta}}} |\boldsymbol{F}(\boldsymbol{\vartheta}^{(i)}; s_j)|\,. \tag{11}$$

We remark that the covariance matrix must be symmetric and positive definite. In order to ensure that it is positive definite we have to handle the case where we try to pick a sensor location twice as in this case the positive definiteness is violated. We do so by setting the argument of the exponential to 10^{-7}. This form of the correlation error reduces the utility when sensors are placed too close together and prevents excessive clustering of the sensors [65,66].

We wish to identify the locations $\boldsymbol{s}$ yielding the largest information gain about the unknown parameter $\boldsymbol{\vartheta}$ of the disturbance. A measure for information gain is defined through the Kullback-Leibler (KL) divergence between the prior belief of the parameter values and the posterior belief, i.e. after

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 October 2019 doi:10.20944/preprints201910.0266.v1

Peer-reviewed version available at Biomimetics 2020, 5, 10; doi:10.3390/biomimetics5010010

measuring the environment. The prior and posterior beliefs are represented through the density functions $p(\boldsymbol{\vartheta})$ and $p(\boldsymbol{\vartheta}|\boldsymbol{y},\boldsymbol{s})$, respectively. We denote by $\mathcal{T}$ the support of $p(\boldsymbol{\vartheta})$. The two densities are connected through Bayes' theorem,

$$p(\boldsymbol{\vartheta}|\boldsymbol{y},\boldsymbol{s}) = \frac{p(\boldsymbol{y}|\boldsymbol{\vartheta},\boldsymbol{s})\, p(\boldsymbol{\vartheta})}{p(\boldsymbol{y})}\,, \tag{12}$$

where $p(\boldsymbol{y}|\boldsymbol{\vartheta},\boldsymbol{s})$ is the likelihood function defined in Eq. (9) and $p(\boldsymbol{y})$ is the normalization constant.

The *utility* function is defined as [67],

$$u(\boldsymbol{s},\boldsymbol{y}) := D_{\mathrm{KL}}(p(\boldsymbol{\vartheta}|\boldsymbol{y},\boldsymbol{s})||p(\boldsymbol{\vartheta})) = \int_{\mathcal{T}} \ln \frac{p(\boldsymbol{\vartheta}|\boldsymbol{y},\boldsymbol{s})}{p(\boldsymbol{\vartheta})}\; p(\boldsymbol{\vartheta}|\boldsymbol{y},\boldsymbol{s})\, \mathrm{d}\boldsymbol{\vartheta}\,. \tag{13}$$

The *expected utility* is defined as the average value over all possible measurements,

$$\begin{aligned} U(\boldsymbol{s}) :&= \mathbb{E}_{\boldsymbol{y}|\boldsymbol{s}}\Big[u(\boldsymbol{s},\boldsymbol{y})\Big] = \int_{\mathcal{Y}} u(\boldsymbol{s},\boldsymbol{y})\; p(\boldsymbol{y}|\boldsymbol{s})\, \mathrm{d}\boldsymbol{y} \\ &= \int_{\mathcal{Y}} \int_{\mathcal{T}} \ln \frac{p(\boldsymbol{\vartheta}|\boldsymbol{y},\boldsymbol{s})}{p(\boldsymbol{\vartheta})}\; p(\boldsymbol{\vartheta}|\boldsymbol{y},\boldsymbol{s})\, \mathrm{d}\boldsymbol{\vartheta}\; p(\boldsymbol{y}|\boldsymbol{s})\, \mathrm{d}\boldsymbol{y}\,, \end{aligned} \tag{14}$$

where $\mathcal{Y}$ is the support of measurements. Using Eq. (12) the expected utility can be expressed as,

$$U(\boldsymbol{s}) = \int_{\mathcal{Y}} \int_{\mathcal{T}} \ln \frac{p(\boldsymbol{y}|\boldsymbol{\vartheta},\boldsymbol{s})}{p(\boldsymbol{y}|\boldsymbol{s})}\; p(\boldsymbol{y}|\boldsymbol{\vartheta},\boldsymbol{s})\; p(\boldsymbol{\vartheta})\, \mathrm{d}\boldsymbol{\vartheta}\, \mathrm{d}\boldsymbol{y}\,. \tag{15}$$

2.2.2. Estimated expected utility for continuous random variables: school location

When $\boldsymbol{\vartheta} = \mathbf{r}$, the random variable takes values in a continuous domain. The estimator for the expected utility in this case can be obtained by approximating the two integrals by Monte Carlo integration using $N_{\boldsymbol{\vartheta}}$ samples from $p(\boldsymbol{\vartheta})$ and $N_{\boldsymbol{y}}$ from $p(\boldsymbol{y}|\boldsymbol{\vartheta},\boldsymbol{s})$ [68]. The resulting estimator is given by,

$$U(\boldsymbol{s}) \approx \hat{U}(\boldsymbol{s}) = \frac{1}{N_{\boldsymbol{\vartheta}} N_{\boldsymbol{y}}} \sum_{j=1}^{N_{\boldsymbol{y}}} \sum_{i=1}^{N_{\boldsymbol{\vartheta}}} \left[\ln p(\boldsymbol{y}^{(i,j)}|\boldsymbol{\vartheta}^{(i)},\boldsymbol{s}) - \ln \left(\frac{1}{N_{\boldsymbol{\vartheta}}} \sum_{k=1}^{N_{\boldsymbol{\vartheta}}} p(\boldsymbol{y}^{(i,j)}|\boldsymbol{\vartheta}^{(k)},\boldsymbol{s}) \right) \right]\,, \tag{16}$$

where $\boldsymbol{\vartheta}^{(i)} \sim p_{\boldsymbol{\vartheta}}$ for $i = 1,\ldots,N_{\boldsymbol{\vartheta}}$ and $\boldsymbol{y}^{(i,j)} \sim \mathcal{N}(\boldsymbol{F}(\boldsymbol{\vartheta}^{(i)};\boldsymbol{s}),\Sigma(\boldsymbol{s}))$ and $j = 1,\ldots,N_{\boldsymbol{y}}$. We remark that the computational complexity of this procedure is mainly determined by the number of Navier-Stokes simulations $N_{\boldsymbol{\vartheta}}$. There is no additional computational burden to compute the $N_{\boldsymbol{y}}$ samples following the measurement error model Eq. (8).

2.2.3. Estimated expected utility for discrete random variables: school size

When $\boldsymbol{\vartheta}$ is a discrete random variable with finite support taking values in the set $\{\boldsymbol{\vartheta}_1,\ldots,\boldsymbol{\vartheta}_{N_{\boldsymbol{\vartheta}}}\}$ the expected utility in Eq. (15) is given by,

$$U(\boldsymbol{s}) = \sum_{i=1}^{N_{\boldsymbol{\vartheta}}} p(\boldsymbol{\vartheta}_i) \int_{\mathcal{Y}} \ln \frac{p(\boldsymbol{y}|\boldsymbol{\vartheta}_i,\boldsymbol{s})}{p(\boldsymbol{y}|\boldsymbol{s})}\; p(\boldsymbol{y}|\boldsymbol{\vartheta}_i,\boldsymbol{s})\;\, \mathrm{d}\boldsymbol{y}\,. \tag{17}$$

Here, $\boldsymbol{\vartheta} = n_f$ represents the number of fish in the leading group. An estimator of the given utility can be obtained by Monte Carlo integration using $N_{\boldsymbol{y}}$ samples from the likelihood distribution $p(\boldsymbol{y}|\boldsymbol{\vartheta}_i,\boldsymbol{s})$. The estimator is given by

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 October 2019 doi:10.20944/preprints201910.0266.v1

Peer-reviewed version available at Biomimetics 2020, 5, 10; doi:10.3390/biomimetics5010010

$$U(\boldsymbol{s}) \approx \hat{U}(\boldsymbol{s}) = \frac{1}{N_{\boldsymbol{y}}} \sum_{j=1}^{N_{\boldsymbol{y}}} \sum_{i=1}^{N_{\boldsymbol{\vartheta}}} p(\boldsymbol{\vartheta}_i) \left[\ln p(\boldsymbol{y}^{(i,j)}|\boldsymbol{\vartheta}_i, \boldsymbol{s}) - \ln \left(\sum_{k=1}^{N_{\boldsymbol{\vartheta}}} p(\boldsymbol{\vartheta}_k) p(\boldsymbol{y}^{(i,j)}|\boldsymbol{\vartheta}_k, \boldsymbol{s}) \right) \right] . \tag{18}$$

where $\boldsymbol{y}^{(i,j)} \sim p(\boldsymbol{y}|\boldsymbol{\vartheta}_i, \boldsymbol{s})$ for $j = 1, \ldots, N_{\boldsymbol{y}}$. Let $\boldsymbol{\varphi}$ be the random variable representing one of the group configurations. Each group configuration is associated with a unique number $\boldsymbol{\varphi}_{i,j}$ for $j = 1, \ldots, n_i$, where n_i is the total number of configurations containing i fish. With this notation, $\boldsymbol{\varphi}$ takes values in the set $\{\boldsymbol{\varphi}_{i,j} \,|\, j = 1, \ldots, n_i \,, i = 1, \ldots, 8\}$. For an example of different configurations see Appendix A The likelihood can be written in terms of the variable $\boldsymbol{\varphi}$ as

$$p(\boldsymbol{y}|\boldsymbol{\vartheta} = \boldsymbol{\vartheta}_i, \boldsymbol{s}) = p(\boldsymbol{y}|\boldsymbol{\varphi} = \{\boldsymbol{\varphi}_{i,1}, \ldots, \boldsymbol{\varphi}_{i,n_i}\}, \boldsymbol{s}) \,. \tag{19}$$

Using the fact that for $i = 1, \ldots, N_{\boldsymbol{\vartheta}}$,

$$p(\boldsymbol{y}, \boldsymbol{\varphi} = \boldsymbol{\varphi}_{k,\ell}|\boldsymbol{\vartheta} = \boldsymbol{\vartheta}_i, \boldsymbol{s}) = 0, \quad \text{for } k \neq i \,,$$

and

$$p(\boldsymbol{y}|\boldsymbol{\vartheta} = \boldsymbol{\vartheta}_i, \boldsymbol{\varphi} = \boldsymbol{\varphi}_{i,\ell}, \boldsymbol{s}) = p(\boldsymbol{y}|, \boldsymbol{\varphi} = \boldsymbol{\varphi}_{i,\ell}, \boldsymbol{s}), \quad \text{for } \ell = 1, \ldots, n_i \,,$$

and the assumption

$$p(\boldsymbol{\varphi} = \boldsymbol{\varphi}_{i,\ell}|\boldsymbol{\vartheta} = \boldsymbol{\vartheta}_i, \boldsymbol{s}) = \frac{1}{n_i}, \quad \text{for } \ell = 1, \ldots, n_i \,,$$

the likelihood in Eq. (19) is written as,

$$\begin{aligned} p(\boldsymbol{y}|\boldsymbol{\vartheta} = \boldsymbol{\vartheta}_i, \boldsymbol{s}) &= \sum_{k=1}^{N_{\boldsymbol{\vartheta}}} \sum_{\ell=1}^{n_i} p(\boldsymbol{y}, \boldsymbol{\varphi} = \boldsymbol{\varphi}_{k,\ell}|\boldsymbol{\vartheta} = \boldsymbol{\vartheta}_i, \boldsymbol{s}) \\ &= \sum_{\ell=1}^{n_i} p(\boldsymbol{y}, \boldsymbol{\varphi} = \boldsymbol{\varphi}_{i,\ell}|\boldsymbol{\vartheta} = \boldsymbol{\vartheta}_i, \boldsymbol{s}) \\ &= \sum_{\ell=1}^{n_i} p(\boldsymbol{y}|\boldsymbol{\vartheta} = \boldsymbol{\vartheta}_i, \boldsymbol{\varphi} = \boldsymbol{\varphi}_{i,\ell}, \boldsymbol{s}) \, p(\boldsymbol{\varphi} = \boldsymbol{\varphi}_{i,\ell}|\boldsymbol{\vartheta} = \boldsymbol{\vartheta}_i) \\ &= \frac{1}{n_i} \sum_{\ell=1}^{n_i} p(\boldsymbol{y}|\boldsymbol{\varphi} = \boldsymbol{\varphi}_{i,\ell}, \boldsymbol{s}) \,. \end{aligned} \tag{20}$$

Notice that the likelihood is a mixture of distributions with equal weights and that $p(\boldsymbol{y}|\boldsymbol{\varphi} = \boldsymbol{\varphi}_{i,\ell}) = \mathcal{N}(\boldsymbol{y}|\boldsymbol{F}(\boldsymbol{\varphi}_{i,\ell}; \boldsymbol{s}), \Sigma(\boldsymbol{s}))$. In order to draw a sample from the likelihood, first we draw an integer ℓ^* with equal probability from 1 to 8 and them draw $y\boldsymbol{y} \sim p(\boldsymbol{y}|\boldsymbol{\varphi} = \boldsymbol{\varphi}_{i,\ell^*})$.

The final form of the estimator is given by

$$\hat{U}(\boldsymbol{s}) = \frac{1}{N_{\boldsymbol{y}}} \sum_{j=1}^{N_{\boldsymbol{y}}} \sum_{i=1}^{N_{\boldsymbol{\vartheta}}} p(\boldsymbol{\vartheta}_i) \left[\ln \frac{1}{n_i} \sum_{\ell=1}^{n_i} p(\boldsymbol{y}^{(i,j)}|\boldsymbol{\varphi}_{i,\ell}, \boldsymbol{s}) - \ln \left(\frac{1}{n_i} \sum_{k=1}^{N_{\boldsymbol{\vartheta}}} p(\boldsymbol{\vartheta}^{(k)}) \sum_{\ell=1}^{n_i} p(\boldsymbol{y}^{(i,j)}|\boldsymbol{\varphi}_{i,\ell}, \boldsymbol{s}) \right) \right] . \tag{21}$$

2.2.4. Optimization of the expected utility function

In order to determine the optimal sensor arrangement we maximize the utility estimator $\hat{U}(\boldsymbol{s})$ described in Eq. (16). It has been observed that the expected utility for many sensors often exhibit many local optima [65,69]. Heuristic approaches, such as the sequential sensor placement algorithm

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 October 2019 doi:10.20944/preprints201910.0266.v1

Peer-reviewed version available at Biomimetics 2020, 5, 10; doi:10.3390/biomimetics5010010

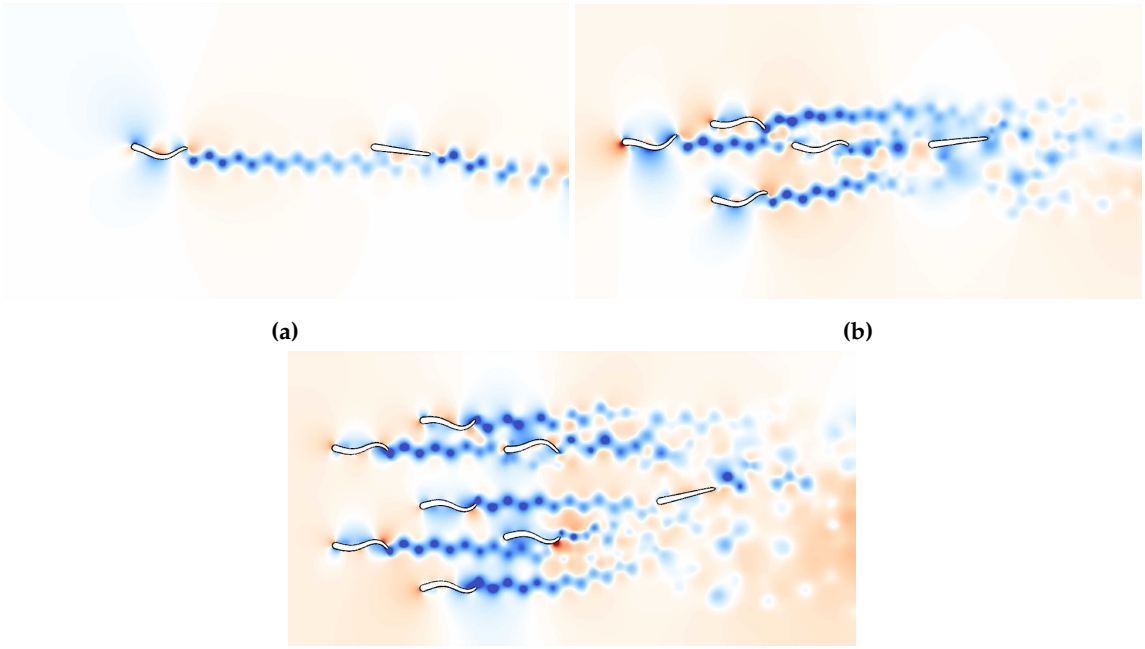


Figure 3. Snapshots of the pressure field in the environment of the coasting swimmer generated by one (Fig. 3a), four (Fig. 3b) and seven (Fig. 3c) schooling swimmers. The snapshots are taken at the moment the measurement was performed for one particular location of the follower in the prior region. High pressure is shown in red and low pressure in blue.

described by [70], have been demonstrated to be effective alternatives. Here, following [70], we perform the optimization iteratively, placing one sensor after the other. We start by placing one sensor $s_1^\star$ by a grid search in the interval $[0, 0.9]$. In the next step we compute the location of the second sensor by setting $\boldsymbol{s} = (s_1^\star, s)$ and repeating the grid search for the new optimal location $s_2^\star$. This procedure is then continued by defining

$$s_i^\star = \underset{s}{\operatorname{argmax}}\ \hat{U}_i(\boldsymbol{s}) \qquad \text{where} \qquad \boldsymbol{s} = (s_1^\star, \ldots, s_{i-1}^\star, s)\,. \tag{22}$$

We note that the scalar variable s denotes the mid-line coordinate of a single sensor-pair, whereas the vector $\boldsymbol{s}$ holds the coordinate of all sensors-pairs. Besides the mentioned advantages, sequential placement allows to quantify the importance of each sensor placed and provides further insight into the resulting distribution of sensors.

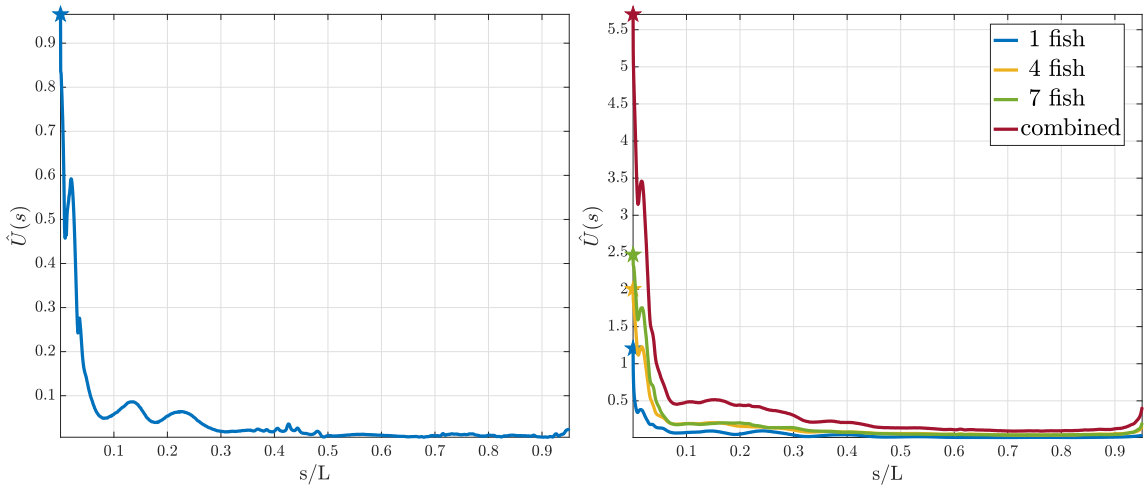
3. Results

We examine the optimal arrangement of pressure gradient and shear stress sensors on the surface of a swimmer trailing a school of self-propelled swimmers. We consider two sensing objectives: (a) the size of the leading school and (b) the relative position of the school. The simulations correspond to a Reynolds number $\mathrm{Re} = \frac{L^2}{\nu} = 2000$. In all experiments we use 4096 points to discretize the horizontal direction $x \in [0, 1]$ and all artificial swimmers have a length of $L = 0.1$.

For the "size of the leading school" experiment the size of the group takes values $\boldsymbol{\vartheta}_i = 1, \ldots, 8$. First we regard one configuration per group-size. In this case inferring the configuration is equivalent to inferring the number of fish in the group. To increase the difficulty we consider n_i different initial configurations. In each configuration we assign a number $\boldsymbol{\varphi}_{i,\ell}$ for $i = 1, \ldots, 8$ and $\ell = 1, \ldots, n_i$. In total, we consider $N_{tot} = \sum_i n_i = 61$ distinct configurations each having the same prior probability $1/N_{tot}$. In Appendix A we present the initial condition for all configurations. The center of mass of the school is located at $x = 0.3$ and in the y-axis in the middle of the vertical extent of the domain. We

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 October 2019 doi:10.20944/preprints201910.0266.v1

Peer-reviewed version available at Biomimetics 2020, 5, 10; doi:10.3390/biomimetics5010010



(a) **(b)**

Figure 4. Utility curves for the first sensor using pressure measurements. In Fig. 4a the utility estimator for the "size of the leading school" experiment is presented. Figure 4b corresponds to the utility estimator for the "relative position" experiment. We show the resulting curves for one, three and seven fish in the leading group and the total expected utility. We observe that although the form does not drastically change, the total utility increases with increasing size of the leading group.

use a controller to fix the distance between x and y coordinates of two fish to $\Delta x = \Delta y = 0.15$, see Section 2.1.1.

For the "relative position" experiment we consider three independent experiments with one, four and seven leading fish. Snapshots of the pressure field for these simulations are presented in Fig. 3. The prior probability for the position of the group is uniform in the domain $[0.6, 0.8] \times [0.1, 0.4]$. The support of the prior probability is discretized with 21×31 gridpoints. Since the experiments are independent, the total expected utility function for the three cases is the sum of the expected utility of each experiment [49].

For both experiments we record the pressure gradient and shear stress on the surface of the fish using the methods discussed in Section 2.1.2. The motion of the fish introduces disturbances on it's own surface. In order to distinguish the self-induced from the environment disturbances we freeze the movement of the following fish and set its curvature to zero. The freezing time is selected by evolving the simulation until the wakes of the leading group are sufficiently mixed and passed the following swimmer. We found that this is the case for $T = 22$. The transition from swimming to coasting motion takes place during the time interval $[T, T+1]$. Finally, we record the pressure gradient and the shear stress at time $T+2$. The resulting sensor-signal associated to the midline coordinates $\boldsymbol{s}$ for a given configuration $\boldsymbol{\vartheta}$ is denoted $\boldsymbol{F}(\boldsymbol{\vartheta}; \boldsymbol{s})$, see Eq. (8).

3.1. Utility function for the first sensor

In this section we discuss the optimal location of a single pressure gradient sensor using the estimators in Eq. (16) and Eq. (21). Higher values of the expected utility correspond to preferable locations for the sensor since the information gain at these sensors is higher. The resulting utilities are plotted in Fig. 4. For all experiments we find that the tip of the head ($s = 0$) exhibits the largest utility independent of the number of fish in the leading group.

At the tip of the head the two symmetrically placed sensors have the smallest distance. In Eq. (10) we have assumed that the two fish-halves are symmetric and uncorrelated. Due to the small distance of the sensors at the head, spatial correlation between the sensors across the fish-halves would decrease the utility of this location. In order to test whether the utility for sensors at the head is influenced by this symmetry assumption, we perform experiments where we place a single sensor on one side of the fish. Again in this case the location at the head is found to have the highest expected utility.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 October 2019 doi:10.20944/preprints201910.0266.v1

Peer-reviewed version available at Biomimetics 2020, 5, 10; doi:10.3390/biomimetics5010010

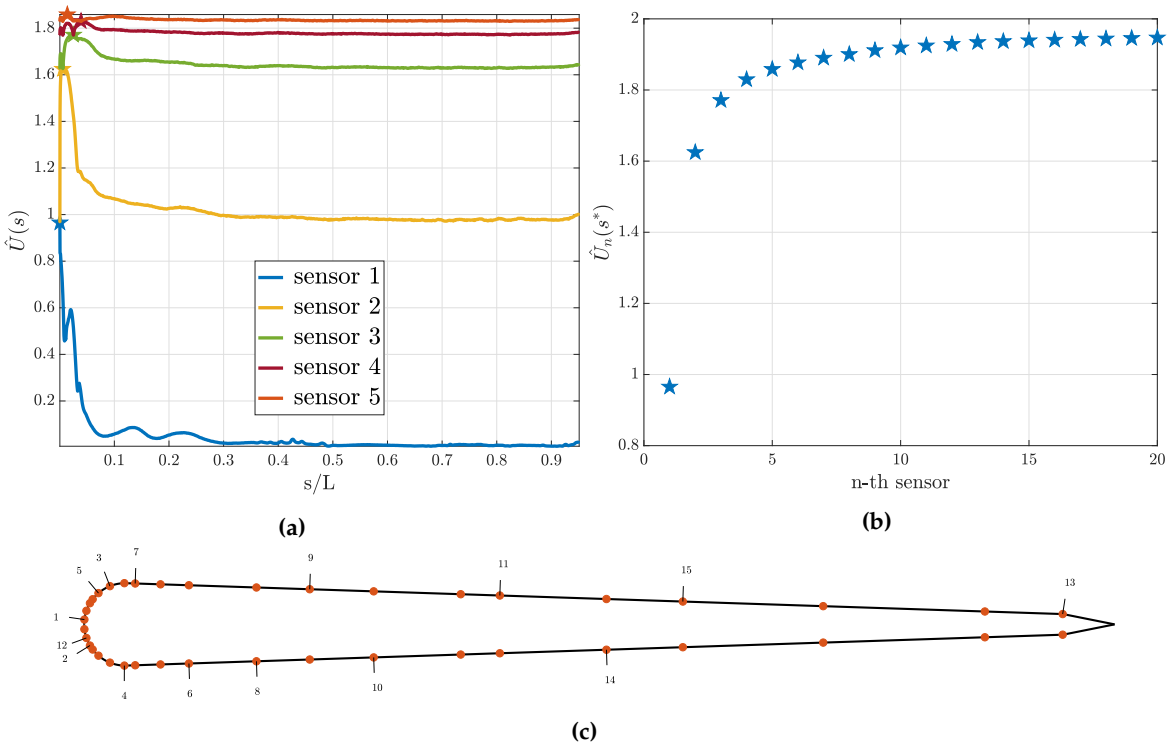


Figure 5. Optimal sensor placement for the pressure sensors and the "size of the leading school" experiment. In Fig. 6a the utility estimator for the first five sensors and in Fig. 6b the value of the utility estimator at the optimal sensor location for the first 20 sensors are presented. In Fig. 6c the distribution of the sensors on the fish surface is presented.

Based on the evidence that the head experiences the largest variance of pressure gradients $\boldsymbol{F}(\boldsymbol{\vartheta}; \boldsymbol{s})$ and that this is correlated with the density of sub-canal neuromast [43], we examine the variance of the values obtained from our solution of the Navier-Stokes equation. We confirm that this observation is consistent with our simulations. We find that independent of the number of fish, the variance in the sensor signal $\mathrm{var}_{\boldsymbol{\vartheta}}(\boldsymbol{F}(\boldsymbol{\vartheta}; \boldsymbol{s}))$ is largest at $s = 0$.

3.2. Sequential sensor placement

In this section we discuss the results of the sequential sensor placement described in Section 2.2.4. For the "size of the leading school" experiment we present the results in Fig. 5. In Fig. 5a the utility curve for the first five sensors is shown. We observe that the utility curve becomes flatter as the number of sensors increase. Furthermore, we observe that the location where the previous sensor was placed is a minimum for the utility for the next sensor. Figure 5b shows the utility estimator at the optimal sensor for up to 20 sensors and it is evident that the value of the expected utility reaches a plateau. In Fig. 5c the location of the sensors on the skin if the fish is presented. The numbers correspond to the iteration in the sequential procedure that the sensor was placed. Note that the sensors are being placed symmetrically.

The optimal sensor placement results for the "relative position" experiment can be found in Fig. 6. Similar to the other experiment the utility curves become flatter after every placed sensor and the location for the previous sensor is a minimum for the utility for the next sensor (see Fig. 6a). We plot the maximum of the utility for up to 20 sensors (see Fig. 6b) and observe a convergence to a constant value. In Fig. 6c the location of the first 20 sensors is presented.

For both experiments it is evident that the utility of the optimal sensor location approaches a constant value. This fact can be explained by recalling that the expected utility Eq. (15) is a measure of the averaged distance between the prior and the posterior distribution. Increasing the number of sensors leads to an increase in the number of measurements. By the Bayesian central limit theorem,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 October 2019 doi:10.20944/preprints201910.0266.v1

Peer-reviewed version available at Biomimetics 2020, 5, 10; doi:10.3390/biomimetics5010010

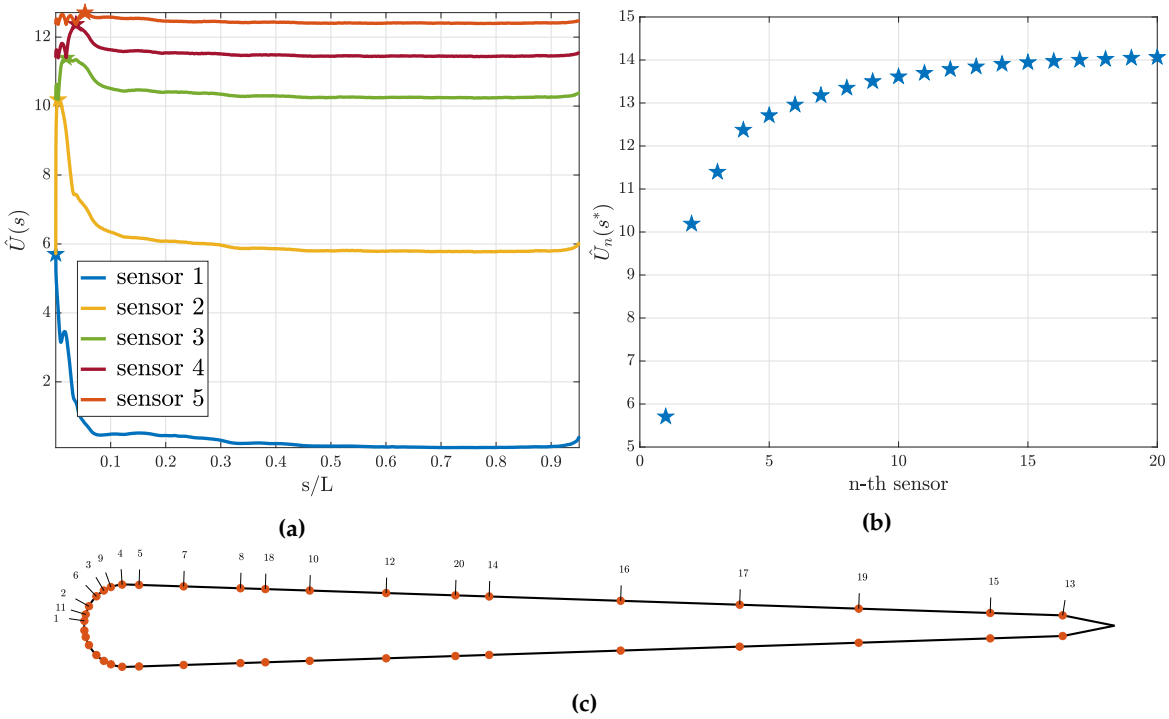


Figure 6. Optimal sensor placement for the pressure gradient sensors for the "relative position" experiment. In Fig. 6a the utility estimator for the first five sensors and in Fig. 6b the value of the utility estimator at the optimal sensor location for the first 20 sensors are presented. In Fig. 6c the distribution of the sensors on the fish surface is presented.

increasing the number of measurements leads to convergence of the posterior to a Dirac distribution. As soon as the posterior has converged, the expected distance from the prior, and thus the expected utility, remains constant.

3.3. Inference of the environment

In this section we demonstrate the importance of the optimal sensor locations and examine the convergence of the posterior distribution. We compute the posterior distribution via Bayes rule given in Eq. (12). We set $\boldsymbol{y} = \mathbf{F}(\boldsymbol{\vartheta}, \boldsymbol{s})$ and compute the posterior for different values of $\boldsymbol{\vartheta}$ in the prior region. We consider measurements collected at: (a) the optimal and (b) the worst sensors location.

The posterior probability for the "size of the leading school" experiment is shown in Fig. 7. We observe that the worst sensor location implies an almost uniform posterior distribution, reflecting that measurements at this sensor carry no information. On the other hand, the posterior distribution for the optimal sensor is more informative. We observe that for groups with small size the follower is able to identify the size with more confidence, as opposed to larger groups. We compare the posterior for an experiment with only one configuration per group-size to an experiment with multiple configurations. For multiple configurations the posterior is less informative. This indicates that the second case occurs to be a more difficult problem. Finally, notice that the posterior for one configuration is symmetric, where when adding multiple configurations this symmetry is broken. This fact is discussed in Appendix B.

The posterior density for the "relative position" with one leading fish is presented in Fig. 8. The posterior for the configuration with three and seven fish is similar. We compute the posterior for measurements at the best and the worst location for one and three sensors. For the three sensors the worst location has been selected in all three phases of the sequential placement. The results for the normalized densities are shown in Fig. 8. We observe that one sensor at the optimal location gives a very peaked posterior. Three optimal sensors can infer the location with low uncertainty. This is not

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 October 2019 doi:10.20944/preprints201910.0266.v1

Peer-reviewed version available at Biomimetics 2020, 5, 10; doi:10.3390/biomimetics5010010

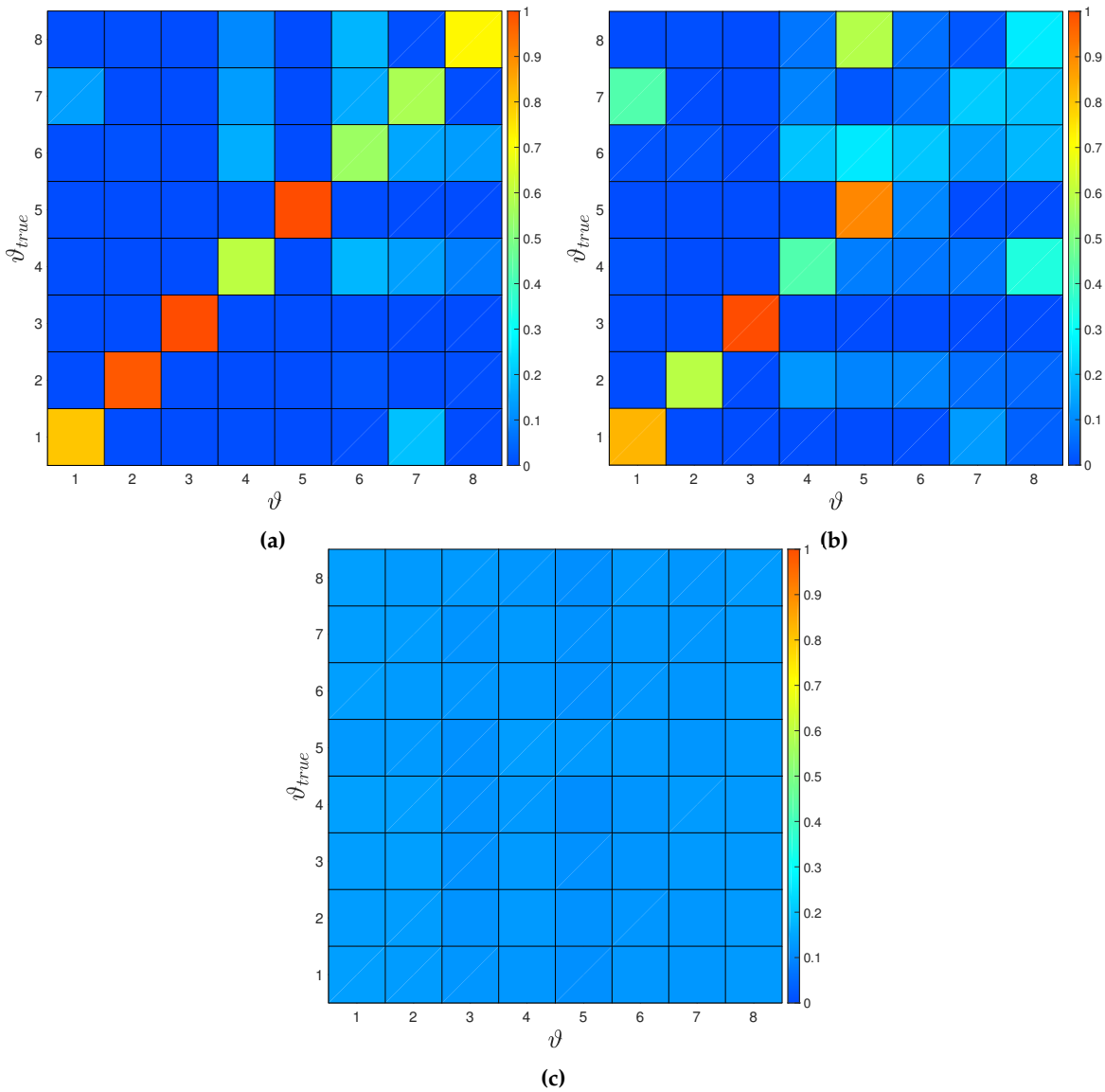
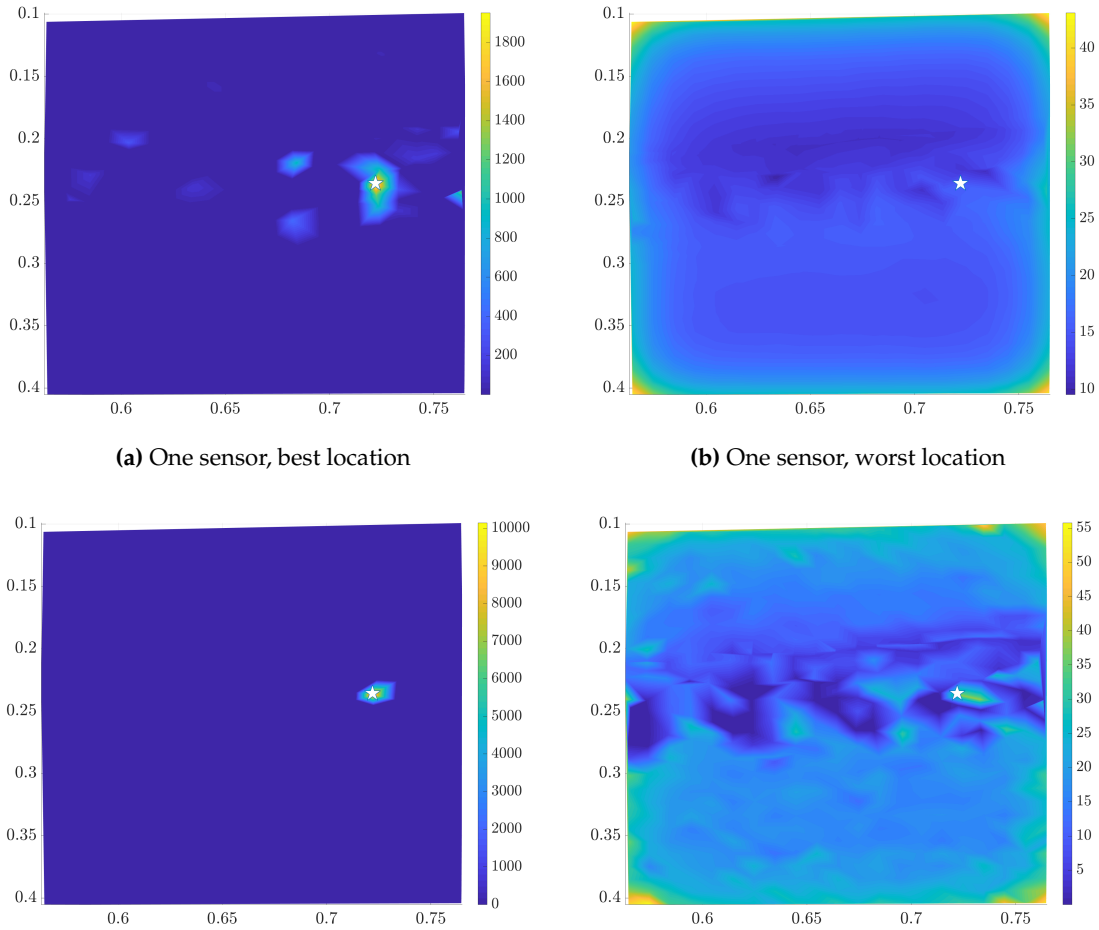


Figure 7. (a) Estimated posterior probability for a single sensor optimally placed and a single configuration per group size. The posterior shows clear peaks at the correct number of fish for all cases, leading to perfect inference of the parameter of inference. The posterior probability for (b) optimal and (c) worst sensor location for multiple configurations per group size. Here, for the optimal sensor location for one, two, three and five fish shows a clear peak for the true size of the group. For the worst sensor location the posterior is almost uniform and does not allow to extract any information about the size of group.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 October 2019 doi:10.20944/preprints201910.0266.v1

Peer-reviewed version available at Biomimetics 2020, 5, 10; doi:10.3390/biomimetics5010010



(a) One sensor, best location

(b) One sensor, worst location

(c) Three sensors, best location

(d) Three sensors, worst location

Figure 8. Estimated posterior for the final location for the best (left column) and worst (right column) sensor-location for one (upper row) and three sensors (lower row). Light colors correspond to high probability density values. We marked the actual location with a white star.

the case for the worst sensors, where adding more sensors does not immediately lead to uncertainty reduction.

3.4. Shear stress sensors

In this section we discuss the results for the optimal positioning of shear stress sensors. We follow the same procedure as in Section 3.1 and Section 3.2. Here, we omit the presentation of the results and focus on the similarities and differences to the pressure gradient sensors.

The optimal location for a single sensor for the "size of the leading school" experiment is at $s^* = 3.01 \cdot 10^{-4}$. For the "relative position" experiment we find the optimal location $s^* = 3.84 \cdot 10^{-4}$. In contrast to the optimal location for one pressure gradient sensor, the found sensor is not at the tip of the head and is at different positions for the two experiments. Examining the variance in the shear signal shows quantitatively the same behaviour as the utility. Comparing the location of the maxima shows that they do not coincide for shear sensors.

We perform sequential placement of 20 sensors. The resulting distribution of sensors is shown in Fig. 9. In Section 3.2 we argue that the expected utility must reach a plateau when placing many sensors using the Bayesian central limit theorem. For shear stress sensors we observe that the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 October 2019 doi:10.20944/preprints201910.0266.v1

Peer-reviewed version available at Biomimetics 2020, 5, 10; doi:10.3390/biomimetics5010010

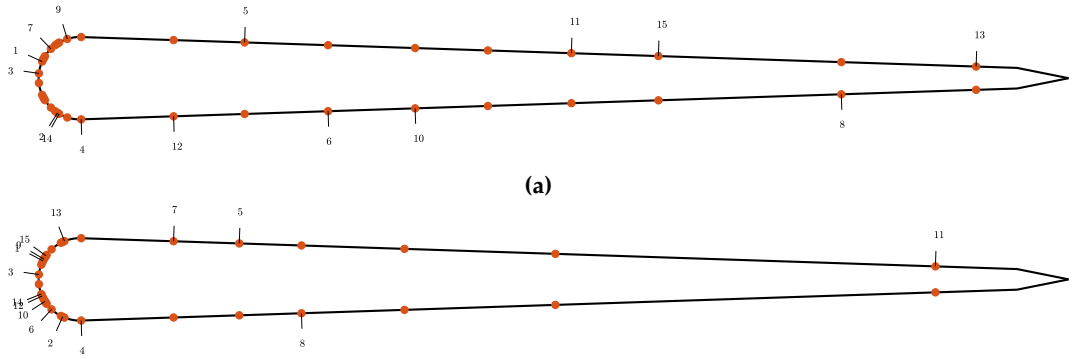


Figure 9. Optimal sensor locations for the shear stress measurements for the "size of the leading school" in Fig. 9a experiment and "relative position" experiment in Fig. 9b.

convergence is slower compared to the pressure gradient sensors. We conclude that the information gain per shear stress sensor placed is lower as for the pressure gradient sensors.

The posterior density obtained for both experiments is less informative when using the same number of sensors. This indicates that shear is a less informative quantity yielding a slower convergence of the posterior.

4. Discussion

We present a study of the optimal sensor locations on a self-propelled swimmer for detecting the number and location of a leading group of swimmers. This optimization combines Bayesian experimental design with large scale simulations of the two dimensional Navier-Stokes equations. Mimicking the function of sensory organs in real fish, we used the shear stress and pressure gradient on the surface of the fish to determine the sensor feedback generated by a disturbance in the flow field.

The optimization was performed for different configurations of fish, ranging from a simple leader-follower configuration with two fish, to a group of up to eight fish leading a single follower. We regarded two types of information, the number of fish in the leading group and the relative location of the leading group. We find that although the general shape of the utility function varies between the two objectives, the preferred location of the first sensor on the head of the fish is consistent. Furthermore, we find that the objective is only weakly influenced when varying the number of members in the leading group.

We perform a sequential sensor placement and find that the utility converges to a constant value and thus we can conclude that few sensors suffice to infer the quantities of the surrounding flow. Indeed we find that the optimal sensor locations correspond to a posterior distribution that is strongly peaked around the true value of the quantity of interest.

In summary, we find that changing the number of fish in the leading group does not influence the followers ability to infer the location of the leading group, for groups with small size. Furthermore we were able to show that choosing the locations for the measurements in a systematic way we are able to infer the number of fish in the leading group and the location of our agent to high accuracy. Ongoing work examines the inclusion of further parameters such as the motion of the swimmers in optimally detecting disturbances from neighboring fish.

Author Contributions: Conceptualization, Georgios Arampatzis and Petros Koumoutsakos; Data curation, Pascal Weber; Formal analysis, Pascal Weber and Georgios Arampatzis; Funding acquisition, Petros Koumoutsakos; Investigation, Pascal Weber; Methodology, Georgios Arampatzis, Siddhartha Verma and Costas Papadimitriou; Project administration, Petros Koumoutsakos; Resources, Petros Koumoutsakos; Software, Pascal Weber and Guido Novati; Supervision, Georgios Arampatzis and Petros Koumoutsakos; Validation, Pascal Weber and Guido Novati; Visualization, Pascal Weber and Georgios Arampatzis; Writing – original draft, Pascal Weber; Writing – review & editing, Pascal Weber, Georgios Arampatzis, Guido Novati, Siddhartha Verma, Costas Papadimitriou and Petros Koumoutsakos.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 October 2019 doi:10.20944/preprints201910.0266.v1

Peer-reviewed version available at Biomimetics 2020, 5, 10; doi:10.3390/biomimetics5010010

Funding: We would like to acknowledge the computational time at Swiss National Supercomputing Center (CSCS) under the project s929. We gratefully acknowledge support from the European Research Council (ERC) Advanced Investigator Award (No. 341117).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A Configurations

The configuration used for the number of fish experiment. For the configurations with three rows the vertical extent $y \in [0, 0.5]$ was discretized using 2048 gridpoints, for the ones with four rows it was extended to $y \in [0, 0.75]$ and discretized using 3072 gridpoints.

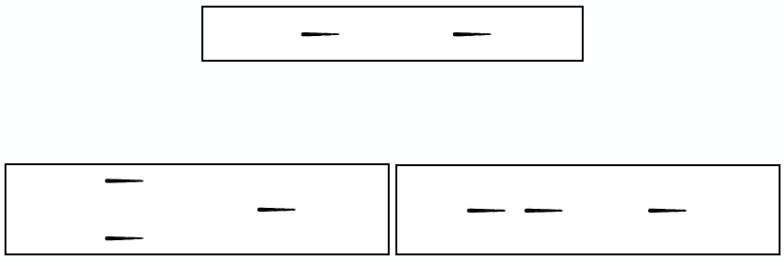


Figure A1. Configurations for two fish.

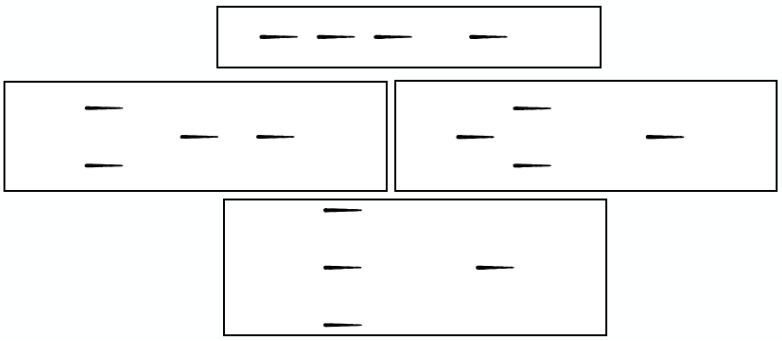


Figure A2. Configurations for three fish.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 October 2019 doi:10.20944/preprints201910.0266.v1

Peer-reviewed version available at Biomimetics 2020, 5, 10; doi:10.3390/biomimetics5010010

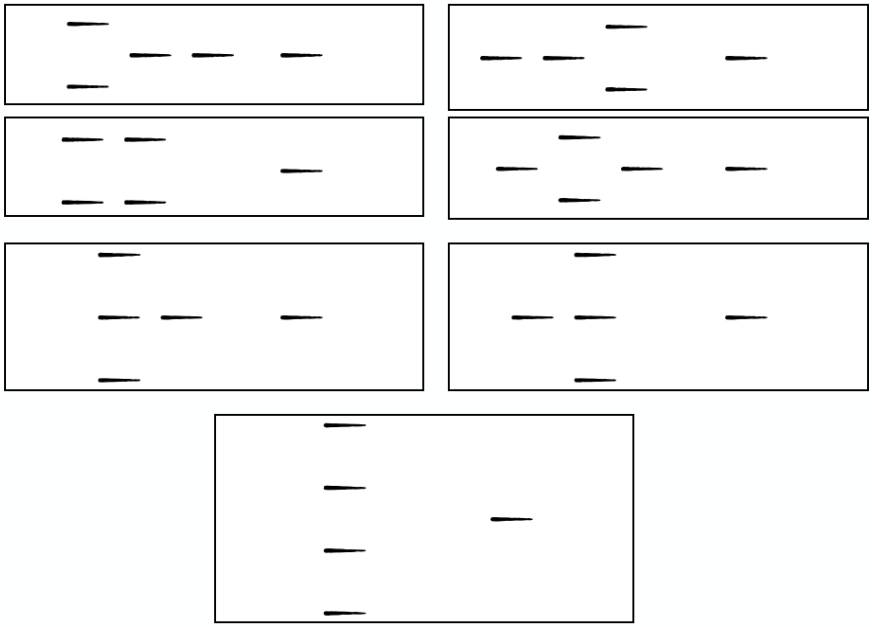


Figure A3. Configurations for four fish.

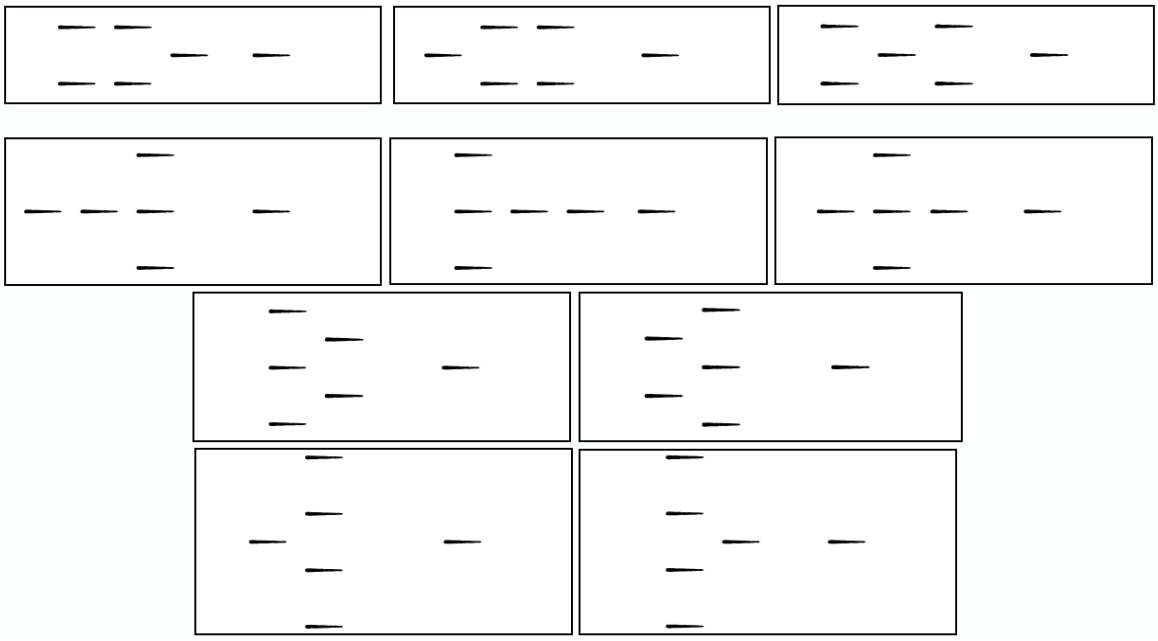


Figure A4. Configurations for five fish.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 October 2019 doi:10.20944/preprints201910.0266.v1

Peer-reviewed version available at Biomimetics 2020, 5, 10; doi:10.3390/biomimetics5010010

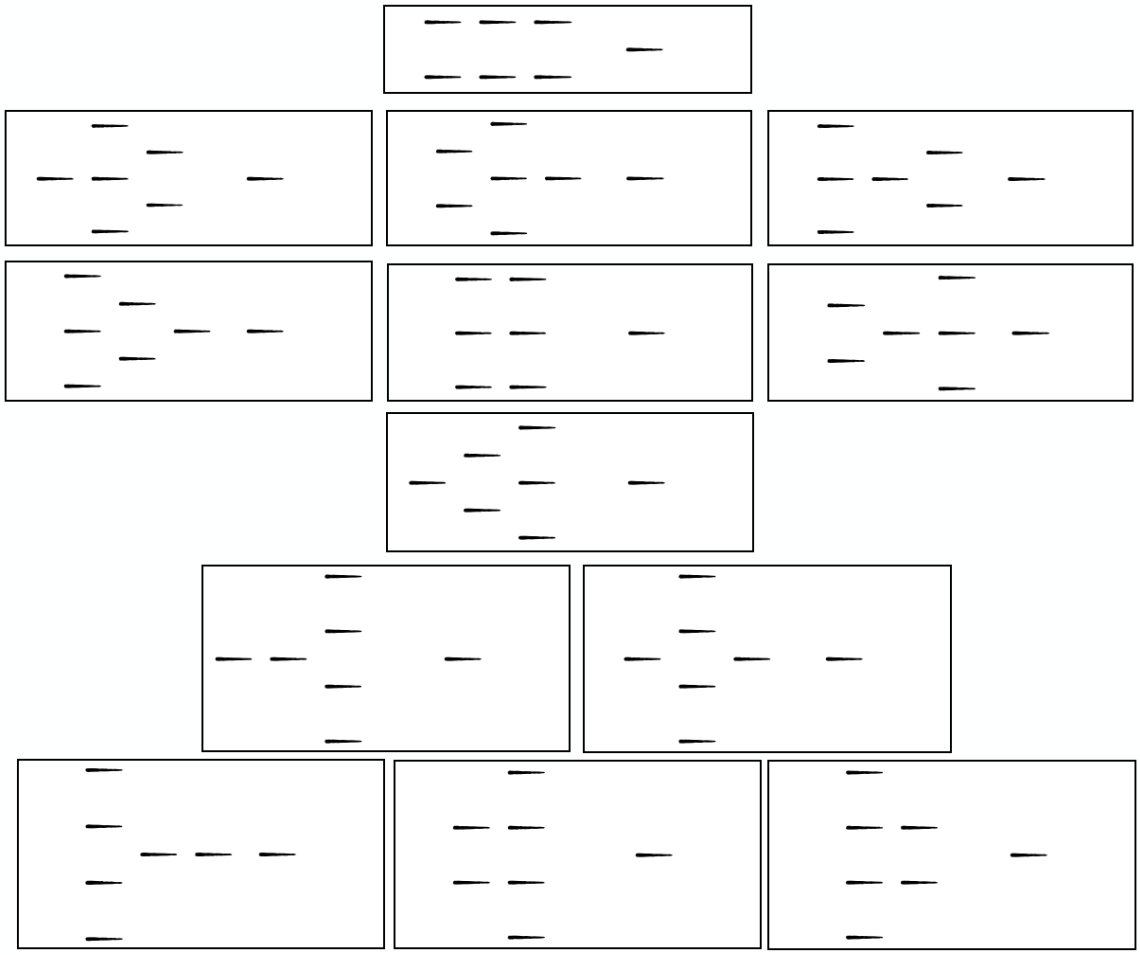


Figure A5. Configurations for six fish.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 October 2019 doi:10.20944/preprints201910.0266.v1

Peer-reviewed version available at Biomimetics 2020, 5, 10; doi:10.3390/biomimetics5010010

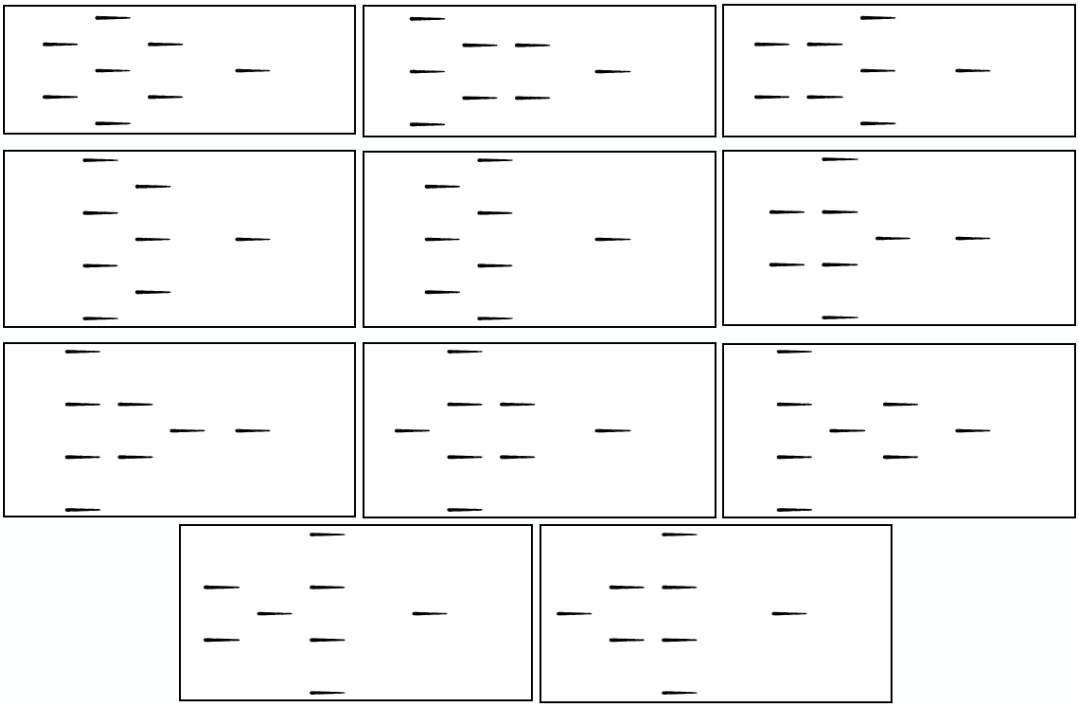


Figure A6. Configurations for seven fish.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 October 2019 doi:10.20944/preprints201910.0266.v1

Peer-reviewed version available at Biomimetics 2020, 5, 10; doi:10.3390/biomimetics5010010

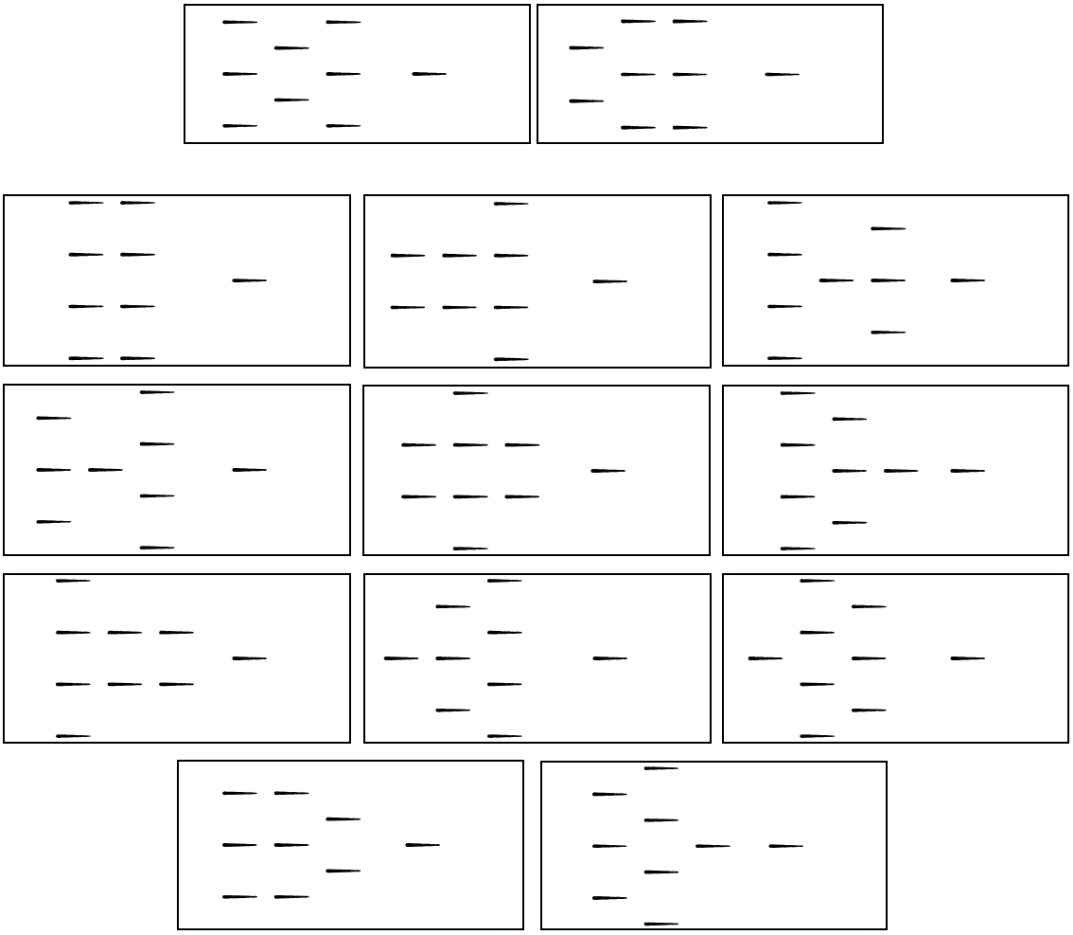


Figure A7. Configurations for eight fish.

Appendix B Posterior is not symmetric

The posterior in Fig. 7 has no mirror symmetry along the diagonal axis. This observation indicates that the posterior is not symmetric with respect to an exchange of the parameter $\boldsymbol{\vartheta}$ we try to infer and the parameter used in the simulation $\boldsymbol{\vartheta}_{\text{true}}$. In the following we want to formally examine this observation. For sake of brevity we neglect the influence on the sensor location $\boldsymbol{s}$. The posterior for some $\boldsymbol{\vartheta}_i$ with $i = 1, \ldots, 8$ is given by Bayes rule

$$p(\boldsymbol{\vartheta}_i|\boldsymbol{y}) \propto p(\boldsymbol{\vartheta}_i)p(\boldsymbol{y}|\boldsymbol{\vartheta}_i)\,. \tag{A1}$$

In Section 2.2.3 we showed that for multiple configurations $\boldsymbol{\varphi}_{i,\ell}$ for a given number of fish $\boldsymbol{\vartheta}_i$ the likelihood is an equally weighted mixture of Gaussian distributions. Further assuming an uniform prior distribution, we have

$$p(\boldsymbol{\vartheta}_i)p(\boldsymbol{y}|\boldsymbol{\vartheta}_i) = \frac{1}{8}\frac{1}{n_i}\sum_{\ell=1}^{n_i}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(\boldsymbol{y}-\mathbf{F}(\boldsymbol{\varphi}_{i,\ell}))^2}{2\sigma^2}\right)\,. \tag{A2}$$

Let us take $\boldsymbol{y} = \mathbf{F}(\boldsymbol{\varphi}^*)$ for some fixed configuration $\boldsymbol{\varphi}^*$ belonging to some number of fish $\boldsymbol{\vartheta}^*$. If we replace $\boldsymbol{\varphi}^* \to \boldsymbol{\varphi}_{i,\ell^*}$ for some fixed configuration $\boldsymbol{\varphi}_{i,\ell^*}$ corresponding to $\boldsymbol{\vartheta}_i$ and $\boldsymbol{\varphi}_{i,\ell} \to \boldsymbol{\varphi}_{*,\ell}$ for the number of fish $\boldsymbol{\vartheta}^*$, we realize that for $n_i, n_* \neq 1$ the exponents in the resulting sum of Gaussian densities are different

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 October 2019 doi:10.20944/preprints201910.0266.v1

Peer-reviewed version available at Biomimetics 2020, 5, 10; doi:10.3390/biomimetics5010010

$$\frac{1}{8}\frac{1}{n_i}\sum_{\ell=1}^{n_i}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(\mathbf{F}(\boldsymbol{\varphi}^*)-\mathbf{F}(\boldsymbol{\varphi}_{i,\ell}))^2}{2\sigma^2}\right)\neq\frac{1}{8}\frac{1}{n_*}\sum_{\ell'=1}^{n_*}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(\mathbf{F}(\boldsymbol{\varphi}_{i,\ell^*})-\mathbf{F}(\boldsymbol{\varphi}_{*,\ell'}))^2}{2\sigma^2}\right), \tag{A3}$$

We remark that in case we have only one configuration $n_i = 1$ for all $i = 1, \ldots, 8$ the contrary is true and the posterior becomes symmetric with respect to an exchange of the parameter $\boldsymbol{\vartheta}$ we try to infer and the parameter used in the simulation $\boldsymbol{\vartheta}_{\text{true}}$.

1. Ward, A.J.W.; Sumpter, D.J.T.; Couzin, I.D.; Hart, P.J.B.; Krause, J. Quorum decision-making facilitates information transfer in fish shoals. *Proceedings of the National Academy of Sciences* **2008**, *105*, 6948–6953, [https://www.pnas.org/content/105/19/6948.full.pdf]. doi:10.1073/pnas.0710344105.
2. Puckett, J.G.; Pokhrel, A.R.; Giannini, J.A. Collective gradient sensing in fish schools. *Scientific Reports* **2018**, *8*, 7587. doi:10.1038/s41598-018-26037-9.
3. Dykgraaf, S. Untersuchungen über die Funktion der Seitenorgane an Fischen. *Zeitschrift für vergleichende Physiologie* **1933**, *20*, 162–214. doi:10.1007/BF00340757.
4. Dykgraaf, S. The functioning and significance of the lateral-line organs. *Biological reviews of the Cambridge Philosophical Society* **1963**, *38*, 51–105. doi:10.1111/j.1469-185X.1963.tb00654.x.
5. Bleckmann, H.; Przybilla, A.; Klein, A.; Schmitz, A.; Kunze, S.; Brücker, C., Station Holding of Trout: Behavior, Physiology and Hydrodynamics. In *Nature-Inspired Fluid Mechanics: Results of the DFG Priority Programme 1207 "Nature-inspired Fluid Mechanics" 2006-2012*; Springer Berlin Heidelberg: Berlin, Heidelberg, 2012; pp. 161–177. doi:10.1007/978-3-642-28302-4_10.
6. Sutterlin, A.; Waddy, S. Possible Role of the Posterior Lateral Line in Obstacle Entrainment by Brook Trout (Salvelinus fontinalis). *Journal of the Fisheries Research Board of Canada* **2011**, *32*, 2441–2446. doi:10.1139/f75-281.
7. Akanyeti, O.; Venturelli, R.; Visentin, F.; Chambers, L.; Megill, W.M.; Fiorini, P. What information do Kármán streets offer to flow sensing? *Bioinspiration & Biomimetics* **2011**, *6*, 036001. doi:10.1088/1748-3182/6/3/036001.
8. Chambers, L.D.; Akanyeti, O.; Venturelli, R.; Ježov, J.; Brown, J.; Kruusmaa, M.; Fiorini, P.; Megill, W.M. A fish perspective: detecting flow features while moving using an artificial lateral line in steady and unsteady flow. *J. R. Soc. Interface* **2014**, *11*. doi:10.1098/rsif.2014.0467.
9. von Campenhausen, C.; Riess, I.; Weissert, R. Detection of stationary objects by the blind Cave FishAnoptichthys jordani (Characidae). *Journal of comparative physiology* **1981**, *143*, 369–374. doi:10.1007/BF00611175.
10. Hassan, E.S. Hydrodynamic Imaging of the Surroundings by the Lateral Line of the Blind Cave Fish Anoptichthys jordani. The Mechanosensory Lateral Line; Coombs, S.; Görner, P.; Münz, H., Eds.; Springer New York: New York, NY, 1989; pp. 217–227.
11. Windsor, S.P.; Norris, S.E.; Cameron, S.M.; Mallinson, G.D.; Montgomery, J.C. The flow fields involved in hydrodynamic imaging by blind Mexican cave fish (Astyanax fasciatus). Part I: open water and heading towards a wall. *Journal of Experimental Biology* **2010**, *213*, 3819–3831, [https://jeb.biologists.org/content/213/22/3819.full.pdf]. doi:10.1242/jeb.040741.
12. Windsor, S.P.; Norris, S.E.; Cameron, S.M.; Mallinson, G.D.; Montgomery, J.C. The flow fields involved in hydrodynamic imaging by blind Mexican cave fish (Astyanax fasciatus). Part II: gliding parallel to a wall. *Journal of Experimental Biology* **2010**, *213*, 3832–3842, [https://jeb.biologists.org/content/213/22/3832.full.pdf]. doi:10.1242/jeb.040790.
13. Hoekstra, D.; Janssen, J. Non-visual feeding behavior of the mottled sculpin, Cottus bairdi, in Lake Michigan. *Environmental Biology of Fishes* **1985**, *12*, 111–117. doi:10.1007/BF00002763.
14. Pitcher, T.; Partridge, B.; Wardle, C. A blind fish can school. *Science* **1976**, *194*, 963–965, [https://science.sciencemag.org/content/194/4268/963.full.pdf]. doi:10.1126/science.982056.
15. Satou, M.; Takeuchi, H.A.; Nishii, J.; Tanabe, M.; Kitamura, S.; Okumoto, N.; Iwata, M. Behavioral and electrophysiological evidences that the lateral line is involved in the inter-sexual vibrational communication

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 October 2019 doi:10.20944/preprints201910.0266.v1

Peer-reviewed version available at Biomimetics 2020, 5, 10; doi:10.3390/biomimetics5010010

of the himé salmon (landlocked red salmon, Oncorhynchus nerka). *Journal of Comparative Physiology A* **1994**, *174*, 539–549. doi:10.1007/BF00217373.

16. Huijbers, C.M.; Nagelkerken, I.; Lössbroek, P.A.C.; Schulten, I.E.; Siegenthaler, A.; Holderied, M.W.; Simpson, S.D. A test of the senses: Fish select novel habitats by responding to multiple cues. *Ecology* **2012**, *93*, 46–55, [https://esajournals.onlinelibrary.wiley.com/doi/pdf/10.1890/10-2236.1]. doi:10.1890/10-2236.1.
17. Montgomery, J.C.; Baker, C.F.; Carton, A.G. The lateral line can mediate rheotaxis in fish. *Nature* **1997**, *389*, 960–963.
18. Coombs, S.; Janssen, J.; Webb, J.F. Diversity of lateral line systems: Evolutionary and functional considerations. In *Sensory biology of aquatic animals*; Springer, 1988; pp. 553–593.
19. Coombs, S.; Görner, P.; Münz, H. A Brief Overview of the Mechanosensory Lateral Line System and the Contributions to This Volume. The Mechanosensory Lateral Line; Coombs, S.; Görner, P.; Münz, H., Eds.; Springer New York: New York, NY, 1989; pp. 3–5.
20. Denton, E.J.; Gray, J.A.B. Some Observations on the Forces Acting on Neuromasts in Fish Lateral Line Canals. The Mechanosensory Lateral Line; Coombs, S.; Görner, P.; Münz, H., Eds.; Springer New York: New York, NY, 1989; pp. 229–246.
21. Coombs, S.; Braun, C.B., Information Processing by the Lateral Line System. In *Sensory Processing in Aquatic Environments*; Springer New York: New York, NY, 2003; pp. 122–138. doi:10.1007/978-0-387-22628-6_7.
22. Coombs, S.; Netten, S.V. The Hydrodynamics and Structural Mechanics of the Lateral Line System. In *Fish Biomechanics*; Academic Press, 2005; Vol. 23, *Fish Physiology*, pp. 103 – 139. doi:https://doi.org/10.1016/S1546-5098(05)23004-2.
23. Bleckmann, H. Peripheral and central processing of lateral line information. *Journal of Comparative Physiology A* **2008**, *194*, 145–158. doi:10.1007/s00359-007-0282-2.
24. Jiang, Y.; Ma, Z.; Zhang, D. Flow field perception based on the fish lateral line system. *Bioinspiration & Biomimetics* **2019**, *14*, 041001. doi:10.1088/1748-3190/ab1a8d.
25. Engelmann, J.; Hanke, W.; Mogdans, J.; Bleckmann, H. Hydrodynamic stimuli and the fish lateral line. *Nature* **2000**, *408*, 1476–4687. doi:10.1038/35040706.
26. Kottapalli, A.G.P.; Asadnia, M.; Miao, J.M.; Barbastathis, G.; Triantafyllou, M.S. A flexible liquid crystal polymer MEMS pressure sensor array for fish-like underwater sensing. *Smart Mater. Struct.* **2012**, *21*, 115030.
27. Tao, J.; Yu, X. Hair flow sensors: from bio-inspiration to bio-mimicking - A review. *Smart Mater. Struct.* **2012**, *21*, 113001.
28. Asadnia, M.; Kottapalli, A.G.P.; Miao, J.; Warkiani, M.E.; Triantafyllou, M.S. Artificial fish skin of self-powered micro-electromechanical systems hair cells for sensing hydrodynamic flow phenomena. *J. R. Soc. Interface* **2015**, *12*. doi:10.1098/rsif.2015.0322.
29. Kottapalli, A.G.P.; Bora, M.; Sengupta, D.; Miao, J.; Triantafyllou, M.S. Hydrogel-CNT Biomimetic Cilia for Flow Sensing. 2018 IEEE SENSORS, 2018, pp. 1–4. doi:10.1109/ICSENS.2018.8589917.
30. Yang, Y.; Chen, J.; Engel, J.; Pandya, S.; Chen, N.; Tucker, C.; Coombs, S.; Jones, D.L.; Liu, C. Distant touch hydrodynamic imaging with an artificial lateral line. *Proceedings of the National Academy of Sciences* **2006**, *103*, 18891–18895, [https://www.pnas.org/content/103/50/18891.full.pdf]. doi:10.1073/pnas.0609274103.
31. Yang, Y.; Nguyen, N.; Chen, N.; Lockwood, M.; Tucker, C.; Hu, H.; Bleckmann, H.; Liu, C.; Jones, D.L. Artificial lateral line with biomimetic neuromasts to emulate fish sensing. *Bioinspir. Biomim.* **2010**, *5*, 016001.
32. Kruusmaa, M.; Fiorini, P.; Megill, W.; de Vittorio, M.; Akanyeti, O.; Visentin, F.; Chambers, L.; El Daou, H.; Fiazza, M.; Ježov, J.; Listak, M.; Rossi, L.; Salumae, T.; Toming, G.; Venturelli, R.; Jung, D.S.; Brown, J.; Rizzi, F.; Qualtieri, A.; Maud, J.L.; Liszewski, A. FILOSE for Svenning: A Flow Sensing Bioinspired Robot. *IEEE Robot. Autom. Mag.* **2014**, *21*, 51–62. doi:10.1109/MRA.2014.2322287.
33. DeVries, L.; Lagor, F.D.; Lei, H.; Tan, X.; Paley, D.A. Distributed flow estimation and closed-loop control of an underwater vehicle with a multi-modal artificial lateral line. *Bioinspir. Biomim.* **2015**, *10*, 025002.
34. Strokina, N.; Kämäräinen, J.; Tuhtan, J.A.; Fuentes-Pérez, J.F.; Kruusmaa, M. Joint Estimation of Bulk Flow Velocity and Angle Using a Lateral Line Probe. *IEEE Trans. Instrum. Meas.* **2016**, *65*, 601–613. doi:10.1109/TIM.2015.2499019.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 October 2019 doi:10.20944/preprints201910.0266.v1

Peer-reviewed version available at Biomimetics 2020, 5, 10; doi:10.3390/biomimetics5010010

35. Triantafyllou, M.S.; Weymouth, G.D.; Miao, J. Biomimetic Survival Hydrodynamics and Flow Sensing. *Annual Review of Fluid Mechanics* **2016**, *48*, 1–24, [https://doi.org/10.1146/annurev-fluid-122414-034329]. doi:10.1146/annurev-fluid-122414-034329.
36. Xu, Y.; Mohseni, K. A Pressure Sensory System Inspired by the Fish Lateral Line: Hydrodynamic Force Estimation and Wall Detection. *IEEE Journal of Oceanic Engineering* **2017**, *42*, 532–543. doi:10.1109/JOE.2016.2613440.
37. Sengupta, D.; Chen, S.H.; Kottapalli, A.G.P., Nature-Inspired Self-Powered Sensors and Energy Harvesters. In *Self-Powered and Soft Polymer MEMS/NEMS Devices*; Springer International Publishing: Cham, 2019; pp. 61–81. doi:10.1007/978-3-030-05554-7_3.
38. Zhang, X.; Shan, X.; Shen, Z.; Xie, T.; Miao, J. A New Self-Powered Sensor Using the Radial Field Piezoelectric Diaphragm in d33 Mode for Detecting Underwater Disturbances. *Sensors* **2019**, *19*, 962. doi:10.3390/s19040962.
39. Ježov, J.; Akanyeti, O.; Chambers, L.D.; Kruusmaa, M. Sensing oscillations in unsteady flow for better robotic swimming efficiency. 2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2012, pp. 91–96. doi:10.1109/ICSMC.2012.6377682.
40. Yen, W.; Sierra, D.M.; Guo, J. Controlling a Robotic Fish to Swim Along a Wall Using Hydrodynamic Pressure Feedback. *IEEE Journal of Oceanic Engineering* **2018**, *43*, 369–380. doi:10.1109/JOE.2017.2785698.
41. Krieg, M.; Nelson, K.; Mohseni, K. Distributed sensing for fluid disturbance compensation and motion control of intelligent robots. *Nature machine intelligence* **2019**, *1*. doi:10.1038/s42256-019-0044-1.
42. Ćurčić-Blake, B.; van Netten, S.M. Source location encoding in the fish lateral line canal. *Journal of Experimental Biology* **2006**, *209*, 1548–1559, [https://jeb.biologists.org/content/209/8/1548.full.pdf]. doi:10.1242/jeb.02140.
43. Ristroph, L.; Liao, J.C.; Zhang, J. Lateral Line Layout Correlates with the Differential Hydrodynamic Pressure on Swimming Fish. *Phys. Rev. Lett.* **2015**, *114*, 018102. doi:10.1103/PhysRevLett.114.018102.
44. Zhang, F.; Lagor, F.; Yeo, D.; Washington, P.; Paley, D. Distributed Flow Sensing Using Bayesian Estimation for a Flexible Fish Robot. Proceedings of the ASME 2015 Dynamic Systems and Control Conference, 2015. doi:10.1115/DSCC2015-9732.
45. Ahrari, A.; Lei, H.; Sharif, M.A.; Deb, K.; Tan, X. Design optimization of artificial lateral line system under uncertain conditions. 2015 IEEE Congress on Evolutionary Computation (CEC), 2015, pp. 1807–1814. doi:10.1109/CEC.2015.7257106.
46. Ahrari, A.; Lei, H.; Sharif, M.A.; Deb, K.; Tan, X. Reliable underwater dipole source characterization in 3D space by an optimally designed artificial lateral line system. *Bioinspir. Biomim.* **2017**, *12*, 036010.
47. Colvert, B.; Alsalman, M.; Kanso, E. Classifying vortex wakes using neural networks. *Bioinspiration & Biomimetics* **2018**, *13*, 025003. doi:10.1088/1748-3190/aaa787.
48. Xu, D.; Lv, Z.; Zeng, H.; Bessaih, H.; Sun, B. Sensor placement optimization in the artificial lateral line using optimal weight analysis combining feature distance and variance evaluation. *ISA Transactions* **2019**, *86*, 110 – 121. doi:https://doi.org/10.1016/j.isatra.2018.10.039.
49. Verma, S.; Papadimitriou, C.; Lüthen, N.; Arampatzis, G.; Koumoutsakos, P. Optimal sensor placement for artificial swimmers. Under consideration for publication in J. Fluid Mech.
50. Kern, S.; Koumoutsakos, P. Simulations of optimized anguilliform swimming. *The Journal of experimental biology* **2006**, *209*, 4841–4857. doi:10.1242/jeb.02526.
51. Gazzola, M.; Chatelain, P.; van Rees, W.M.; Koumoutsakos, P. Simulations of single and multiple swimmers with non-divergence free deforming geometries. *J. Comput. Phys.* **2011**, *230*, 7093–7114. doi:10.1016/j.jcp.2011.04.025.
52. Kern, S.; Chatelain, P.; Koumoutsakos, P. Modeling, Simulation and Optimization of Anguilliform Swimmers. *Bio-mechanisms of Swimming and Flying: Fluid Dynamics, Biomimetic Robots, and Sports Science* **2008**, p. 167.
53. Carling, J.; Williams, T.L.; Bowtell, G. Self-propelled anguilliform swimming: Simultaneous solution of the two-dimensional Navier-Stokes equations and Newton's laws of motion. *Journal of Experimental Biology* **1998**, *201*, 3143–3166.
54. Angot, P.; Bruneau, C.H.; Fabrie, P. A penalization method to take into account obstacles in incompressible viscous flows. *Numerische Mathematik* **1999**, *81*, 497–520.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 October 2019 doi:10.20944/preprints201910.0266.v1

Peer-reviewed version available at Biomimetics 2020, 5, 10; doi:10.3390/biomimetics5010010

55. Coquerelle, M.; Cottet, G.H. A vortex level set method for the two-way coupling of an incompressible fluid with colliding rigid bodies. *J. Comput. Phys.* **2008**, *227*, 9121–9137.
56. Gazzola, M.; Vasilyev, O.V.; Koumoutsakos, P. Shape optimization for drag reduction in linked bodies using evolution strategies. *COMPUTERS & STRUCTURES* **2011**, *89*, 1224–1231. doi:10.1016/j.compstruc.2010.09.001.
57. Towers, J.D. Finite difference methods for approximating Heaviside functions. *Journal of Computational Physics* **2009**, *228*, 3478–3489. doi:10.1016/j.jcp.2009.01.026.
58. Chorin, A.J. Numerical Solution of the Navier-Stokes Equations. *Mathematics of Computation* **1968**, *22*, 745–762.
59. Novati, G.; Verma, S.; Alexeev, D.; Rossinelli, D.; van Rees, W.M.; Koumoutsakos, P. Synchronised Swimming of Two Fish. *arXiv preprint arXiv:1610.04248* **2016**.
60. Kroese, A.B.; Schellart, N.A. Velocity- and acceleration-sensitive units in the trunk lateral line of the trout. *J. Neurophysiol.* **1992**, *68*, 2212–2221. doi:10.1152/jn.1992.68.6.2212.
61. Bleckmann, H.; Zelick, R. Lateral line system of fish. *Integr. Zool.* **2009**, *4*, 13–25.
62. Jones, E.; Oliphant, T.; Peterson, P.; others. SciPy: Open source scientific tools for Python, 2001–. [Online; accessed <today>].
63. Dierckx, P. An algorithm for smoothing, differentiation and integration of experimental data using spline functions. *Journal of Computational and Applied Mathematics* **1975**, *1*, 165 – 184. doi:https://doi.org/10.1016/0771-050X(75)90034-0.
64. Huan, X.; Marzouk, Y.M. Simulation-based optimal Bayesian experimental design for nonlinear systems. *Journal of Computational Physics* **2013**, *232*, 288 – 317. doi:https://doi.org/10.1016/j.jcp.2012.08.013.
65. Papadimitriou, C.; Lombaert, G. The effect of prediction error correlation on optimal sensor placement in structural dynamics. *Mech. Syst. Signal Process.* **2012**, *28*, 105–127.
66. Simoen, E.; Papadimitriou, C.; Lombaert, G. On prediction error correlation in Bayesian model updating. *J. Sound Vib.* **2013**, *332*, 4136–4152.
67. Ryan, K.J. Estimating Expected Information Gains for Experimental Designs With Application to the Random Fatigue-Limit Model. *J. Comput. Graph. Stat.* **2003**, *12*, 585–603. doi:10.1198/1061860032012.
68. Huan, X.; Marzouk, Y.M. Simulation-based optimal Bayesian experimental design for nonlinear systems. *J. Comput. Phys.* **2013**, *232*, 288–317.
69. Papadimitriou, D.I.; Papadimitriou, C. Optimal sensor placement for the estimation of turbulence model parameters in CFD. *International Journal for Uncertainty Quantification* **2015**, *5*, 545–568.
70. Papadimitriou, C. Optimal sensor placement methodology for parametric identification of structural systems. *Journal of Sound and Vibration* **2004**, *278*, 923–947. doi:10.1016/j.jsv.2003.10.063.