Robust, distributed estimation of internal wave parameters via inter-drogue measurements

Michael Ouimet Jorge Cortés

Abstract

Internal waves are important to oceanographers because, as they travel, they are capable of displacing mass, such as plankton and small fish. This paper considers a group of drogues estimating the physical parameters that determine the dynamics of an ocean linear internal wave. While underwater, individual drogues do not have access to absolute position information and can only rely on inter-drogue measurements. Building on this data and the knowledge of the drogue dynamics under the flow induced by the internal wave, we propose the `Vanishing Distance Derivative Detection Strategy` to allow individual drogues to determine the wave parameters. We analyze the correctness and robustness of this strategy under noiseless and noisy measurements, respectively. We also introduce a general methodology, termed p`th-Order Parameter Fusion`, for combining parameter estimates obtained at different times and characterize the resulting error. Several simulations illustrate our results.

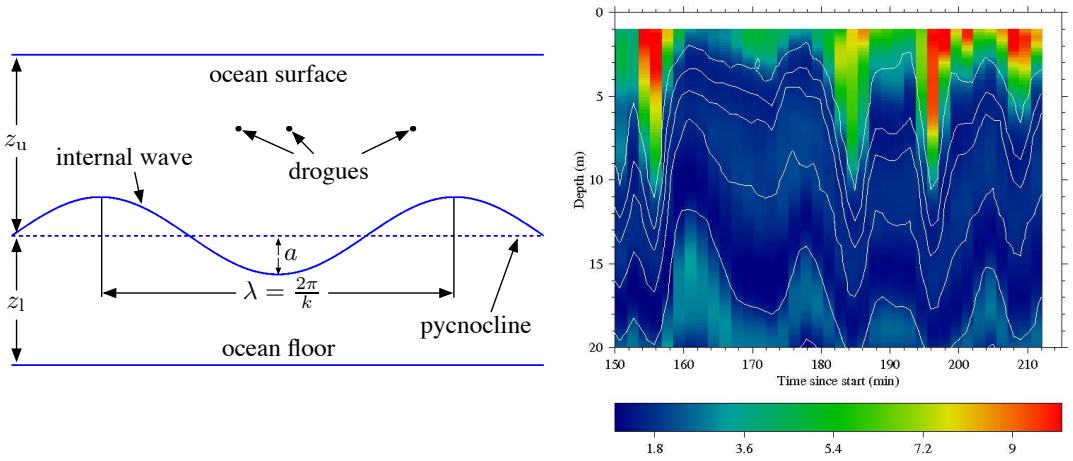
I. Introduction

Internal waves are waves that propagate within a fluid, rather than on its surface. They correspond to moving sinusoidal oscillations in the boundary surface between two layers of a stratified fluid. Of particular relevance are internal waves that move along ocean pycnoclines, which are surfaces of constant density where the vertical rate of change in density is largest. Such internal waves are biologically important as they transport plankton and other organisms and create dense phytoplankton blooms. Because pycnoclines are typically deep below the ocean surface, it is difficult to collect data about internal waves.

A preliminary version of this paper was presented at the 2012 IEEE Conference on Decision and Control, Maui, Hawaii.

The authors are with the Department of Mechanical and Aerospace Engineering, University of California, San Diego, CA 92093, USA, {`miouimet,cortes`}`@ucsd.edu`

This paper tackles this problem using a group of drogues drifting underwater near the internal wave's interface to determine the physical parameters that define its motion. A drogue is a Lagrangian drifter capable of actuating its depth by changing its buoyancy. While underwater, drogues are subject to the flow induced by the motion of the internal wave and do not have access to exact location information. Figure 1 presents a pictorial illustration of the problem setup. The basic premise of the paper is that drogues should be able to extract the information contained in the evolution of the inter-drogue distance and distance derivative measurements to characterize the internal wave. To the authors' knowledge, there is no algorithmic procedure available in the literature to solve this problem.



(a) Schematic of drogues and internal wave (b) Plankton patchiness induced by internal wave

Fig. 1. For a horizontally propagating ocean internal wave, (a) shows its spatial structure at a fixed instant of time whereas (b) shows its temporal structure at a fixed horizontal location. In (a), one can see a vertical cross-section of the ocean perpendicular to the wave propagation direction. A group of drogues float at a constant depth (but not necessarily along a straight line) and do not have access to exact location information. Our objective is to provide drogues with mechanisms that rely only on the relative measurements between them to determine the parameters that uniquely define the internal wave. In (b), one can see data at an anchor station off Mission Beach, CA, taken on April 19, 1997. The plankton's chlorophyll fluorescence (color scale) is depicted as a function of depth and time. Higher fluorescence corresponds to denser patches of plankton in the troughs of a horizontally propagating internal wave. Figure courtesy of Peter Franks, see [1] for additional information.

Literature review: Internal waves are known to be associated with high concentrations of various types of planktonic organisms and small fishes [2], [3], and this makes studying them relevant to oceanographers [4], [1], [5]. Scientists widely use drogues drifting passively as monitoring platforms to gather relevant ocean data [6], [7], [8]. Recent work [9] explores

the possibility of actively selecting tidal currents so that drogues can autonomously reach a desired destination. An increasing body of work in the systems and control literature deals with cooperative networks of agents estimating spatial natural phenomena, including ocean [10], [11], [12], river [13], and hurricane sampling [14]. In these scenarios, agents with limited actuation capabilities are subject to strong flowfields. In the problem considered in this paper, drogues are able to actuate their depth through buoyancy changes, but are completely subject to the force of the internal wave in the flow-wise direction. Because of this, the task of determining the wave parameters can be seen as a data fitting problem [15], [16]. Due to the periodic nature of the inter-drogue distance trajectories, our problem has connections with least-squares spectral analysis problems [17], [18]. In general, however, the fact that the wave parameters appear nonlinearly makes the determination of the exact parameters challenging.

Statement of contributions: We consider the problem of estimating the physical parameters (propagation direction, horizontal wavenumber, frequency, and relative amplitude) of a linear internal wave that is propagating horizontally. A group of underwater drogues with no absolute position information are subject to the flow induced by the internal wave and can only measure inter-drogue distances and distance derivatives. Because the drogues only have access to these relative measurements, they must rely on the presence of other drogues to achieve their task. The benefit obtained here by 'the power of many' in the estimation of the ocean flow field is an original feature of our paper. Our first contribution is the establishment of an analytic expression for the dynamic evolution of the drogues. This expression shows that the motion of each drogue corresponds to a sum of a linear function (which is common to all) and a periodic function in time. In particular, this result implies that the distance function between any two drogues is periodic (with a different period than the internal wave). This analysis sets the basis for our second contribution, which is the design of the `Vanishing Distance Derivative Detection Strategy`. This algorithm builds on the expression for the drogue dynamics and the fact that inter-drogue distance derivatives become close to zero multiple times across a wavelength to estimate the physical parameters of the internal wave. To our knowledge, the proposed algorithm is the first and only method capable of solving the problem formulation described above. We establish the correctness of the `Vanishing Distance Derivative Detection Strategy` in the case where the inter-drogue measurements are noiseless. Specifically, we make precise the range of times along the period of the internal wave when our method can determine

exactly all the parameters. This allows us to give a bound on the minimum required sampling rate. As a third contribution, we characterize the robustness of our strategy by providing explicit bounds of the errors in the parameter estimation as a function of the errors in the acquisition of the inter-drogue measurements. Finally, we develop a general scheme for aggregating the estimates of the parameters provided by our algorithm at different time instants under noisy measurements. Even though we assume the measurement noise to be Gaussian, the highly nonlinear nature of the drogue motion induced by the internal wave makes the distributions of parameters non-Gaussian and therefore, challenging to aggregate. Our aggregation scheme, termed p`th-Order Parameter Fusion`, is based on determining a pth-order approximation of a parameter's distribution. Individual estimates are then fused together assuming the pth-order approximation is exact. The aggregation scheme results in smaller errors than the individual estimates. Several simulations illustrate our correctness and robustness guarantees.

Organization: Section II introduces some basic notation and preliminary notions on derivative estimation. After introducing the internal wave and drogue models, Section III describes the problem statement. Section IV provides the design of the `Vanishing Distance Derivative Detection Strategy` and the analysis guaranteeing its correctness in the absence of noise. Section V establishes the robustness of the algorithm under noise and introduces the p`th-Order Parameter Fusion` to perform the aggregation of estimates. Section VI gathers our conclusions and ideas for future work. Finally, we gather in the appendix a brief exposition on derivative estimation and the proofs of the main results.

II. Preliminaries

Here we present some basic concepts used in the paper, starting with some notational conventions. Let $\mathbb{R}$, $\mathbb{R}_{>0}$, $\mathbb{Z}$, and $\mathbb{Z}_{\geq 1}$, denote the sets of real, positive real, integer, and positive integer numbers, respectively. For $x \in \mathbb{R}$, let $\lfloor x \rfloor \in \mathbb{Z}$ denote the largest integer satisfying $\lfloor x \rfloor \leq x$. For a continuously differentiable function $f : \mathbb{R}^d \to \mathbb{R}$, we let $\partial_k f$ denote the partial derivative with respect to the k-th component. We refer to real-analytic functions simply as 'analytic'. For a vector v, we define the k-th component as $\mathrm{cpnt}_k(\mathbf{v})$.

A reference frame Σ_g in $\mathbb{R}^3$ is composed of an origin $\mathbf{p}_g \in \mathbb{R}^3$ and a set of orthonormal vectors $\{\mathbf{e}_{x_g}, \mathbf{e}_{y_g}, \mathbf{e}_{z_g}\} \subset \mathbb{R}^3$. A point $\mathbf{q}$ and a vector $\mathbf{v}$ can be uniquely expressed with respect to the frame Σ_g and are denoted by $\mathbf{q}^g$ and $\mathbf{v}^g$, respectively. Next, let $\Sigma_b = (\mathbf{p}_b, \{\mathbf{e}_{x_b}, \mathbf{e}_{y_b}, \mathbf{e}_{z_b}\})$

be a reference frame fixed to a moving body. The origin of Σ_b is a point $\mathbf{p}_b$, denoted as $\mathbf{p}_b^g$ when expressed with respect to Σ_g. The orientation of Σ_b is characterized by the rotation matrix Q_b^g whose columns are the vectors $\{\mathbf{e}_{x_b}, \mathbf{e}_{y_b}, \mathbf{e}_{z_b}\}$ expressed with respect to Σ_g. With this notation, a change of reference frame is given by $\mathbf{q}^g = Q_b^g \mathbf{q}^b + \mathbf{p}_b^g$ and $\mathbf{v}^g = Q_b^g \mathbf{v}^b$.

III. Problem statement

This section formulates the problem under study. We begin by presenting the basic model for the motion of a linear internal wave. Then, we describe the capabilities of the group of drogues and discuss the effect that the internal wave has on their dynamics. With these ingredients in place, we formalize the distributed parameter estimation problem.

A. Internal wave model

Let $\Sigma_g = (\mathbf{p}_g, \{\mathbf{e}_{x_g}, \mathbf{e}_{y_g}, \mathbf{e}_{z_g}\})$ be a global reference frame defined as follows: the origin p_g corresponds to an arbitrary point at the surface of the water; the vector $\mathbf{e}_{x_g}$ corresponds to the direction of wave propagation, which is parallel to the ocean bottom, and $\mathbf{e}_{z_g}$ is perpendicular to the ocean bottom, pointing from bottom to surface. For convenience, the coordinates induced by Σ_g are denoted by $\{x, y, z\}$.

As in Figure 1, an internal wave is a wave which travels beneath the surface of the ocean, along a surface of constant water density called pycnocline. We consider an internal wave with amplitude a, frequency ω, propagating horizontally in the x-direction with horizontal wavenumber k, and at the mean depth $-z_{\mathrm{u}}$. The wave depth z_w as a function of x and t is

$$z_w(t, x) = -z_{\mathrm{u}} - a \sin(kx - \omega t + \phi).$$

The parameter ϕ, termed initial phase of the wave, effectively shifts the wave relative to the reference $(x, t) = (0, 0)$. Because of our choice of reference frame, there is no motion in the y-direction. The standard model [4], [1], [19] assumes that vertical velocity varies linearly with depth. This, coupled with the conservation of mass law for an incompressible fluid, gives rise to the following expressions for the horizontal u_{u} and vertical w_{u} velocities of the upper layer,

$$u_{\mathrm{u}}(t, x) = \frac{\omega a}{k z_{\mathrm{u}}} \sin(kx - \omega t + \phi), \tag{1a}$$

$$w_{\mathrm{u}}(t, x, z) = -\frac{z a \omega}{z_{\mathrm{u}}} \cos(kx - \omega t + \phi). \tag{1b}$$

Likewise, the horizontal u_{l} and vertical w_{l} velocities of the lower layer are

$$u_{\text{l}}(t,x) = -\frac{\omega a}{k z_{\text{l}}} \sin(kx - \omega t + \phi),$$
$$w_{\text{l}}(t,x,z) = \frac{z + z_{\text{u}} + z_{\text{l}}}{z_{\text{l}}} a\omega \cos(kx - \omega t + \phi).$$

Assumption 3.1 (Bounds on wave parameters): The linear internal wave model is only accurate for $0 < a/z_{\text{u}} \leq (a/z_{\text{u}})_{\max} = 0.1$. Additionally, the spatial wavelengths of internal waves range from hundreds of meters to tens of kilometers [5]. Since k is inversely proportional to the spatial wavelength, we assume that there exists a bounded interval $[k_{\min}, k_{\max}]$ that k is guaranteed to be in. Finally, since the ratio $\frac{\omega}{k}$ corresponds to the wave's speed, which physically must be bounded, we assume that there exists $\omega_{\max}$ such that $\omega \leq \omega_{\max}$. •

B. Drogue model

A drogue is a submersible buoy which can drift in the ocean, unattached to the ocean floor or a boat, and is able to change its depth in the water by controlling its buoyancy. A drogue can measure the relative distance, the distance derivative, and the orientation to other drogues, as well as depth through sensing (e.g., via acoustic or optical sensors and an onboard compass). However, it does not have access to absolute position because GPS is unavailable underwater.

We consider a group of N drogues. For each $i \in \{1, \ldots, N\}$, let $\Sigma_i = (\mathbf{p}_i, \{\mathbf{e}_{x_i}, \mathbf{e}_{y_i}, \mathbf{e}_{z_i}\})$ be a reference frame fixed to drogue i. The origin $\mathbf{p}_i$ corresponds to the location of the drogue. As in the global coordinate frame Σ_g, $\mathbf{e}_{z_i}$ is perpendicular to the ocean bottom, pointing from bottom to surface. The vectors $\mathbf{e}_{x_i}$ and $\mathbf{e}_{y_i}$ are parallel to the ocean floor, but neither is necessarily oriented in the direction of wave propagation. Thus, each drogue i must determine the angle between $\mathbf{e}_{x_i}$ and $\mathbf{e}_x$, which we denote by θ_i.

Drogues are able to measure inter-drogue distances and distance derivatives. In our treatment, we deal separately with the case of noiseless and noisy measurements. We assume drogues take measurements at a sampling rate of f_s. Thus, at time $t \in \mathbb{R}_{>0}$, drogue i has measurements $\{(\mathbf{d}_{ij}(t_\kappa), \mathbf{d}'_{ij}(t_\kappa))\}_{\kappa \in \{0,\ldots,\lfloor f_s t \rfloor\}}$ at times $t_\kappa = \frac{\kappa}{f_s}$ and for drogues $j \in \{j_1, \ldots, j_M\}$, where these are the M drogues closest to i.

Consider the scenario where drogues move in the upper layer of the internal wave at a constant depth. There is no loss of generality in dealing with this situation, since drogues can control their depth through buoyancy changes. We make the simplifying assumption that the drogue

dynamics under the linear internal wave is Lagrangian. In other words, the dynamics of the drogue position $\mathbf{p} = (p^x, p^y, p^z)$ in the global reference frame is given by

$$\mathbf{p}' = (p'^x, p'^y, p'^z) = (u_{\mathrm{u}}(t, p^x), 0, 0). \tag{2}$$

The absence of motion in the z-direction in this equation is due to the drogue's buoyancy control, which we assume is capable of counteracting the vertical forcing of the internal wave. Since the drogues can measure their depth, we assume there exists an underlying controller which uses these measurements to regulate the drogue at a desired depth. Figure 2 illustrates the time evolution of the x-component of inter-drogue distances as a function of the initial wave phase.

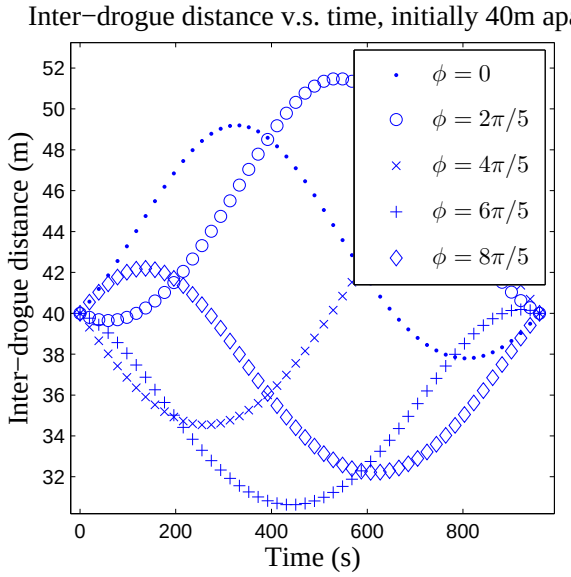


Fig. 2. Inter-drogue distance evolution for drogues initially 40 meters apart, with different initial wave phases.

Remark 3.2 (Kinematic versus dynamical model): The Lagrangian model for motion under the internal wave, cf. (2), is a simplification of the second-order dynamic model, see e.g. [20],

$$mp''^x = -c_d\,|p'^x - u_{\mathrm{u}}(t, p^x)|(p'^x - u_{\mathrm{u}}(t, p^x)), \tag{3a}$$

$$p'^y = 0, \tag{3b}$$

$$mp''^z = -c_d\,|p'^z - w_{\mathrm{u}}(t, p^x, p^z)|(p'^z - w_{\mathrm{u}}(t, p^x, p^z)) + f, \tag{3c}$$

where m denotes the combined drogue mass and inertial added mass [21], c_d is the drag parameter, and f is the buoyancy control input. From this equation, one can derive

$$|u_u(t, p^x(t)) - p'^x(t)| \leq \sqrt{\frac{mu'_{\max}}{c_d}}\tanh\Big(\sqrt{\frac{c_d u'_{\max}}{m}}t + \tanh^{-1}\Big(\sqrt{\frac{c_d}{mu'_{\max}}}u_{\max}\Big)\Big),$$

where $u_{\max} = \frac{\omega a}{k z_u}$ and $u'_{\max} = \frac{\omega a}{k z_u}(k\frac{\omega a}{k z_u} + \omega)$ are bounds on the maximum velocity and acceleration according to the model (1). Following [4], [8], reasonable values for these quantities are $u_{\max} = .02\frac{\text{m}}{\text{s}}$, $u'_{\max} = .00014\frac{\text{m}}{\text{s}^2}$, $m = 1.5$kg, and $c_d = 210\frac{\text{Ns}^2}{m^2}$. For these values, the errors in the drogues' velocities will asymptotically be at most 5% of $u_{\max}$, leading us to favor the kinematic model over the dynamical one. Furthermore, we can see that in the worst case, with $p'^x(0) = 0$ when the drogues are dropped in the water and $u_u(0, p^x(0)) = u_{\max}$, after about 20 seconds, the drogues' velocities will be within 99% of their asymptotic behavior. Thus, after the drogues have been in the water for this long, the drogues' motion can be reasonably well modeled by the kinematic model. This analysis is illustrated in Figure 3. •

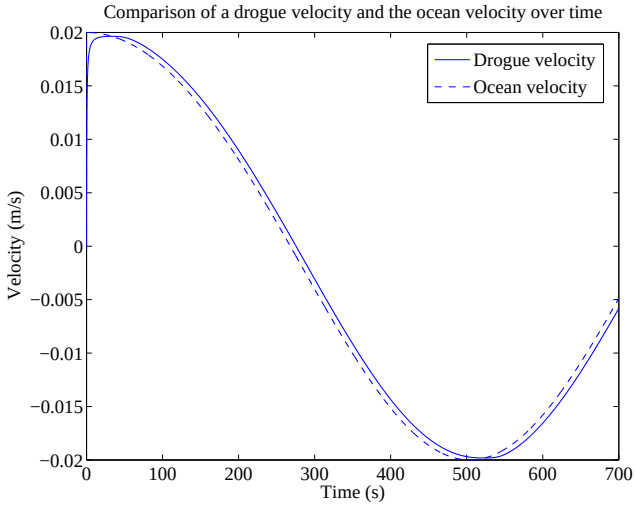


Fig. 3. Comparison of the drogue's velocity as it moves according to the second-order dynamic model (3a) versus according to the kinematic model (2) corresponding to the ocean velocity. Within a short amount of time, the drogue's velocity is very close to the ocean velocity. The wave and drogue parameters are $k = \frac{2\pi}{200}\frac{1}{\text{m}}$, $\omega = \frac{2\pi}{1000}\frac{1}{\text{s}}$, $\frac{a}{z_u} = .1$, $m = 1.5$ kg, and $c_d = 210\frac{\text{Ns}^2}{\text{m}^2}$.

With our models now introduced, we are ready to state the problem of interest in this paper.

Problem statement: A team of N drogues is deployed in the ocean and their motion is governed by an internal wave. Since the drogues may control their depth, assume all are located at the same depth and each one can measure the relative distance, the distance derivative, and orientation to the closest M drogues in their own coordinate frame. The objective is to design a provably correct strategy that allows each drogue i to determine the parameters $\frac{a}{z_u}$, k, ω, and θ_i, defining the motion of the internal wave with the limited information it possesses.

IV. Noise-free parameter estimation

We begin by noting that the dynamic evolution of a drogue under the linear internal wave can be explicitly described in the global reference frame. However, drogues cannot rely on this information as they do not have access to their global coordinates. This motivates our design of methods to determine the wave propagation direction and the internal wave parameters using the distance measurements available to the drogues.

The following result shows that, remarkably, the drogue's dynamic evolution (2) can be described in an analytical way. Its proof is given in the appendix.

Proposition 4.1 (Drogue trajectory): The solution of (2) starting from $\mathbf{p}(0)$ is

$$p^x(t) = \frac{\omega}{k}\Big(1-\sqrt{1-(a/z_{\mathrm{u}})^2}\Big)t + \Xi(t) - \frac{\phi}{k},$$

$$\Xi(t) = \frac{2}{k}\mathrm{atan}\Big(\frac{a}{z_{\mathrm{u}}} - \sqrt{1-\Big(\frac{a}{z_{\mathrm{u}}}\Big)^2}\tan\Big(\frac{\pi t}{T}+\Lambda_0\Big)\Big) - \frac{2\pi}{k}\left\lfloor \frac{t}{T} + \frac{\Lambda_0}{\pi} - \left\lfloor \frac{kp^x(0)+\phi+\pi}{2\pi}\right\rfloor + \frac{1}{2}\right\rfloor + \frac{2\pi}{kT}t,$$

with

$$T = \frac{2\pi}{\omega\sqrt{1-(a/z_{\mathrm{u}})^2}}, \quad \Lambda_0 = \mathrm{atan}\left(\frac{1}{\sqrt{1-(a/z_{\mathrm{u}})^2}}\left(\frac{a}{z_{\mathrm{u}}} - \tan\left(\frac{kp^x(0)+\phi}{2}\right)\right)\right).$$

From Proposition 4.1, we see that the solution of (2) is the sum of a linear function in t and a periodic function Ξ with fundamental period T. Since the linear function does not depend on the initial condition, we deduce that the time evolution of the distance d_{ij}^x between any two drogues i and j is given by (with $\nu = \sqrt{1-(a/z_{\mathrm{u}})^2}$ for brevity)

$$\begin{aligned} d_{ij}^x(t) =& \frac{2}{k}\mathrm{atan}\left(\frac{a}{z_{\mathrm{u}}} - \nu\tan\left(\frac{\nu\omega t}{2}+\Lambda_{0,j}\right)\right) - \frac{2\pi}{k}\left\lfloor \frac{t}{T} + \frac{\Lambda_{0,j}}{\pi} - \left\lfloor \frac{kp_j^x(0)+\phi+\pi}{2\pi}\right\rfloor + \frac{1}{2}\right\rfloor \\ & - \frac{2}{k}\mathrm{atan}\left(\frac{a}{z_{\mathrm{u}}} - \nu\tan\left(\frac{\nu\omega t}{2}+\Lambda_{0,i}\right)\right) + \frac{2\pi}{k}\left\lfloor \frac{t}{T} + \frac{\Lambda_{0,i}}{\pi} - \left\lfloor \frac{kp_i^x(0)+\phi+\pi}{2\pi}\right\rfloor + \frac{1}{2}\right\rfloor \end{aligned}$$

and is periodic with period T (as was numerically observed in Figure 2). However, from a drogue's viewpoint, two facts make this expression impractical: first, drogues do not have access to distances in the global reference frame and, second, since absolute position is not available, they are also unaware of their phase with respect to the internal wave. Even without these two hurdles, the highly nonlinear dependence of this expression on the parameters makes the results for standard least-squares data fitting methods [17], [18] not directly applicable.

These observations motivate the ensuing discussion describing a method to determine the internal wave parameters in the absence of measurement noise. Our treatment is presented for

a generic drogue $i \in \{1, \ldots, N\}$ which requires inter-drogue distance and distance derivative measurements from its nearest 4 neighbors, denoted by $\{j_1, j_2, j_3, j_4\}$. Before getting into the details, we provide a brief overview of the algorithm design.

[Informal description]: Section IV-A describes a method to determine the wave propagation direction. With this information available, drogues can project their inter-drogue measurements along the wave propagation direction. Section IV-B uses knowledge of the drogue's dynamics to find the unmeasurable relative phase between the drogue and the wave as a function of the unknown horizontal wavenumber and the measurable inter-drogue data. Using this function, the algorithm can determine the true value of the horizontal wavenumber employing the data. After this, Section IV-C finds the amplitude ratio and frequency by solving equations derived from the drogue dynamics.

A. Wave propagation direction

Here we describe the method that drogues use to determine the wave propagation direction in their own body coordinates. Recall that the drogues are at the same depth but may be arbitrarily located in the $x-y$ plane. Consider the inter-drogue distance to the jth drogue, $j \in \{j_1, \ldots, j_4\}$, as measured by i in its own body coordinates,

$$\mathbf{d}_{ij} = \mathbf{p}_j - \mathbf{p}_i = (d_{ij}^{x_i}, d_{ij}^{y_i}, 0).$$

Figure 4 depicts the drogue i's own body coordinates, inter-drogue distance measurements, and the direction of wave propagation. For drogues undergoing motion purely caused by an internal

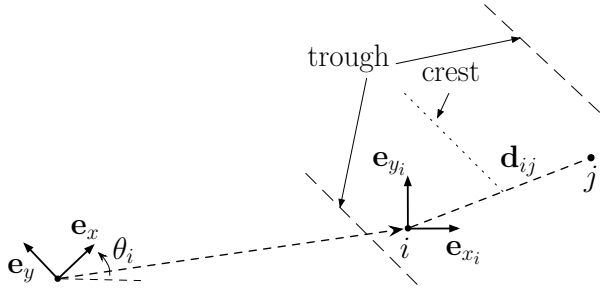


Fig. 4. Illustration of drogue and wave orientations on the drogue's reference frame.

wave, inter-drogue distances in their own body reference frame can be projected onto the global

reference frame $\mathbf{d}_{ij}^g = Q_i^g \mathbf{d}_{ij}$ via the transformation matrix Q_i^g,

$$Q_i^g = \begin{bmatrix} \cos\theta_i & -\sin\theta_i & 0 \\ \sin\theta_i & \cos\theta_i & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The global coordinate frame is useful because the inter-drogue distance in the y-direction is constant, i.e., $d_{ij}^{y\prime} = d_{ij}^{x_i\prime}\sin\theta_i + d_{ij}^{y_i\prime}\cos\theta_i = 0$. Since θ_i is constant, it can easily be found using the measurements available: $\theta_i = \text{atan}\Big(\frac{-d_{ij}^{y_i\prime}}{d_{ij}^{x_i\prime}}\Big)$.

Since both θ_i and $\theta_i + \pi$ fit this relation, we assume the drogues can differentiate the true θ_i. One way to accomplish this is to surface after at least one wave period and use GPS to determine which way the drogue has drifted.

B. Horizontal wavenumber via vanishing derivative

Here we describe a method to determine the horizontal wavenumber of the internal wave. Since the only dynamics are in the x-direction and each drogue i has determined θ_i as described in Section IV-A, from now on, we simply denote d_{ij}^x by d_{ij}. Thus, for each $i \in \{1, \ldots, N\}$, the following dynamics describe the drogues' motion in the x-direction,

$$d_{ij}' = 2\frac{\omega a}{k z_{\text{u}}} \sin\Big(\frac{k d_{ij}}{2}\Big)\cos\Big(\frac{k d_{ij}}{2} + \psi_i\Big), \quad \forall j \in \{j_1, \ldots, j_4\}, \tag{4a}$$

$$\psi_i' = \omega\Big(\frac{a}{z_{\text{u}}}\sin(\psi_i) - 1\Big), \tag{4b}$$

where $\psi_i = k p_i^x - \omega t + \phi$ is the phase of the wave relative to drogue i, which is unmeasurable to it. We note that each inter-drogue distance equation (4a) contains the unknowns ψ_i, $\frac{a}{z_{\text{u}}}$, ω, and k. Our strategy proceeds by deriving an equation which is only a function of measurement data and the horizontal wavenumber, and then determining conditions under which the correct value can be obtained. We later use this knowledge to determine the remaining parameters. In what follows, we make a notational distinction between $\mathcal{k}$ (the horizontal wavenumber interpreted as a variable) and k (the correct horizontal wavenumber that we seek to determine).

First, we show a basic result about the evolution of inter-drogue distances. This helps us formulate assumptions on the initial drogue locations to make the ensuing strategy applicable.

Lemma 4.2 (Inter-drogue distance bound): If at time t_0 the inter-drogue distance between i and j is bounded by $0 < |\frac{k}{2} d_{ij}(t_0)| < \frac{\pi}{2}$, then this bound holds for all $t \geq t_0$.

Proof: We reason by contradiction. Assume there exists $t_* > t_0$ such that $|\frac{k}{2} d_{ij}(t_*)| \in \{0, \frac{\pi}{k}\}$. From (4a), we deduce that $d'_{ij}(t) = 0$ for all $t > t_*$, i.e., the inter-drogue distance stays constant. However, this contradicts the fact that the inter-drogue distances are periodic with period T. ■

We present the assumptions on ordering of inter-drogue distances as well as lower- and upper-bounds on initial inter-drogue distances.

Assumption 4.3 (Inter-drogue distance assumptions): Without loss of generality, we assume $0 < d^x_{ij_1} < d^x_{ij_2} < d^x_{ij_3} < d^x_{ij_4}$. By Lemma 4.2, there exists an $\alpha > 0$ such that each drogue is at least α away from all other drogues in the x-direction, i.e.,

$$d^x_{ij}(t) > \alpha, \quad \forall i \in \{1, \ldots, N\},\ \forall j \in \{j_1, j_2, j_3, j_4\},\ \forall t \geq t_0$$

We assume that initially drogue i and its 4 nearest neighbors are within one spatial wavelength of the internal wave, encapsulated by $d^x_{ij_4}(t_0) < \frac{2\pi}{k_{\max}}$, where k will be in $[k_{\min}, k_{\max}]$ by Assumption 3.1. Furthermore, by Lemma 4.2, this holds $\forall t \geq t_0$. Similarly, we assume that initially

$$d^x_{ij_2}(t_0) < \pi\left(\frac{1}{k_{\max}} - \frac{10}{99 k_{\min}}\right) - \alpha.$$

By using the dynamics of (4a) and Assumption 3.1, this condition ensures that drogue i and its two closest neighbors are always within half a spatial wavelength minus α of each other. •

1) Unmeasurable relative phase: We begin by showing that the unmeasurable relative phase ψ_i can be explicitly expressed in terms of any two inter-drogue distances, which we choose as d_{ij_1} and d_{ij_2}, the distance derivatives d'_{ij_1} and d'_{ij_2}, and $\mathcal{k}$, as the sum of two functions:

$$\psi_i(\mathcal{k}, d_{ij_1}, d_{ij_2}, d'_{ij_1}, d'_{ij_2}) = \nu(\mathcal{k}, d_{ij_1}, d_{ij_2}, d'_{ij_1}, d'_{ij_2}) + \mu(\mathcal{k}, d_{ij_1}, d_{ij_2}, d'_{ij_1}, d'_{ij_2}). \tag{5}$$

The function ν captures the basic structure of ψ_i; it is derived by taking the quotient of two equations in the form of (4a) and solving for ψ_i. One loses information of the sign of d'_{ij_1} and d'_{ij_2} when one takes the ratio of them. Thus, $\{0, -\pi\}$-valued function μ determines which of the two solutions to $\tan(\psi_i) = C$ is the physically meaningful one. Specifically,

$$\nu(\mathcal{k}, d_{ij_1}, d_{ij_2}, d'_{ij_1}, d'_{ij_2}) = \operatorname{atan}\left(\frac{d'_{ij_1} \sin(\frac{\mathcal{k} d_{ij_2}}{2}) \cos(\frac{\mathcal{k} d_{ij_2}}{2}) - d'_{ij_2} \sin(\frac{\mathcal{k} d_{ij_1}}{2}) \cos(\frac{\mathcal{k} d_{ij_1}}{2})}{d'_{ij_1} \sin^2(\frac{\mathcal{k} d_{ij_2}}{2}) - d'_{ij_2} \sin^2(\frac{\mathcal{k} d_{ij_1}}{2})}\right),$$

$$\mu(\mathcal{k}, d_{ij_1}, d_{ij_2}, d'_{ij_1}, d'_{ij_2}) = \begin{cases} 0, & \mathcal{F}(\mathcal{k}, d_{ij_2}, d'_{ij_2}, \nu) > 0 \vee \mathcal{F}(\mathcal{k}, d_{ij_1}, d'_{ij_1}, \nu) > 0, \\ -\pi, & \mathcal{F}(\mathcal{k}, d_{ij_2}, d'_{ij_2}, \nu) < 0 \vee \mathcal{F}(\mathcal{k}, d_{ij_1}, d'_{ij_1}, \nu) < 0, \end{cases}$$

where

$$\mathcal{F}(\mathcal{k}, d_{ij_2}, d'_{ij_2}, \nu) = \frac{d'_{ij_2}}{\cos(\frac{\mathcal{k} d_{ij_2}}{2} + \nu(\mathcal{k}, d_{ij_1}, d_{ij_2}, d'_{ij_1}, d'_{ij_2}))}.$$

Next, we present a result on the properties of ν, which are needed to determine k.

Lemma 4.4 (Smoothness properties of the relative phase): Under Assumption 4.3, the functions ν and $\partial_{\mathcal{k}}\nu$ are Lipschitz continuous with respect to $\frac{d'_{ij_1}}{d'_{ij_2}}$.

Proof: One can show that ν and $\partial_{\mathcal{k}}\nu$ are differentiable with respect to $\frac{d'_{ij_1}}{d'_{ij_2}}$ if

$$0 < \frac{k d_{ij_1}}{2} < \frac{k d_{ij_2}}{2} < \frac{\pi}{2}. \tag{6}$$

By Assumption 4.3, the drogues will be in a compact subset of (6), which shows the result. ■

We refer to the Lipschitz constants for ν and $\partial_{\mathcal{k}}\nu$ as $L_1(\alpha)$ and $L_2(\alpha)$, respectively. In general, ψ_i is a complex trigonometric function of $\mathcal{k}$. However, there are specific time instants for which its expression simplifies considerably, as the following result shows.

Lemma 4.5 (Simplifying expression of the relative phase): For time t_{crit}, let $d'_{ij}(t_{\text{crit}}) = 0$. Assume that drogues i and j are placed so that $0 < |\frac{k}{2} d_{ij}(t_{\text{crit}})| < \frac{\pi}{2}$. Then $\psi_i(t_{\text{crit}}) = \pm\frac{\pi}{2} - \frac{k}{2} d_{ij}(t_{\text{crit}})$.

Proof: The only thing that we need to justify is the existence of t_{crit}. Once this has been established, the explicit expression of $\psi_i(t_{\text{crit}})$ readily follows from (4a). From Proposition 4.1, recall that d_{ij} is bounded and periodic. From (4b), we know that for any $t \in \mathbb{R}_{>0}$, $\psi'_i(t) \leq -\omega(1 - \frac{a}{z_{\text{u}}}) < 0$, where we have used the fact that $\frac{a}{z_{\text{u}}} \leq .1$. Therefore, looking at (4a), one can conclude the existence of t_{crit} within one period T when $d'_{ij}(t_{\text{crit}}) = 0$. ■

Recall that Lemma 4.2 guarantees that the assumptions of Lemma 4.5 are not difficult to ensure.

2) Distance rate quotient: Next, we note that the ratio of inter-drogue distance equations of the form (4a), say d'_{ij_3}/d'_{ij_4}, eliminates ω and a/z_{u}. These observations lead us to define the *distance rate quotient* function as follows. Let

$$\text{dr}(\mathcal{k}, \psi_i, d_{ij}) = \sin\Big(\frac{\mathcal{k} d_{ij}}{2}\Big) \cos\Big(\frac{\mathcal{k} d_{ij}}{2} + \psi_i\Big).$$

Then, define

$$\begin{aligned}\mathrm{drq}(\mathcal{k},\mathbf{D}) &= \frac{\sin(\frac{\mathcal{k}d_{ij_3}}{2})\cos(\frac{\mathcal{k}d_{ij_3}}{2}+\psi_i(\mathcal{k},d_{ij_1},d_{ij_2},d'_{ij_1},d'_{ij_2}))}{\sin(\frac{\mathcal{k}d_{ij_4}}{2})\cos(\frac{\mathcal{k}d_{ij_4}}{2}+\psi_i(\mathcal{k},d_{ij_1},d_{ij_2},d'_{ij_1},d'_{ij_2}))}-\frac{d'_{ij_3}}{d'_{ij_4}}\\ &= \frac{\sin(\frac{\mathcal{k}d_{ij_3}}{2})\cos(\frac{\mathcal{k}d_{ij_3}}{2}+\boldsymbol{\nu}(\mathcal{k},d_{ij_1},d_{ij_2},d'_{ij_1},d'_{ij_2}))}{\sin(\frac{\mathcal{k}d_{ij_4}}{2})\cos(\frac{\mathcal{k}d_{ij_4}}{2}+\boldsymbol{\nu}(\mathcal{k},d_{ij_1},d_{ij_2},d'_{ij_1},d'_{ij_2}))}-\frac{d'_{ij_3}}{d'_{ij_4}}\\ &= \frac{\mathrm{dr}(\mathcal{k},\boldsymbol{\nu}(\mathcal{k},d_{ij_1},d_{ij_2},d'_{ij_1},d'_{ij_2}),d_{ij_3})}{\mathrm{dr}(\mathcal{k},\boldsymbol{\nu}(\mathcal{k},d_{ij_1},d_{ij_2},d'_{ij_1},d'_{ij_2}),d_{ij_4})}-\frac{d'_{ij_3}}{d'_{ij_4}},\end{aligned} \tag{7}$$

where $\mathbf{D}=(d_{ij_1},d_{ij_2},d_{ij_3},d_{ij_4},d'_{ij_1},d'_{ij_2},d'_{ij_3},d'_{ij_4})$ is the collection of all 4 inter-drogue distances and distance derivatives. The second equality comes from noting that drq takes the same value for either value that the function μ takes. By definition, $\mathcal{k}=k$ satisfies

$$\mathrm{drq}(k,\mathbf{D})=0. \tag{8}$$

In principle, there could be additional roots to this equation. This is what we investigate next.

3) Determining the horizontal wavenumber: Our goal now is to determine conditions that guarantee that only k is a solution to (8). The following result precisely characterizes how small the ratio d'_{ij_1}/d'_{ij_2} should be in order to guarantee that k is the unique value that satisfies (8). Its proof is given in the appendix.

Proposition 4.6: (Range of suitable derivative ratios for determining k) Assuming internal wave parameters are within the bounds in Assumption 3.1, consider noiseless inter-drogue distance and distance derivative measurements $\mathbf{D}$ satisfying Assumption 4.3 and $\left|\frac{d'_{ij_1}}{d'_{ij_2}}\right|<\delta(\alpha,k_{\min})$ where $\delta(\alpha,k_{\min})=\min\left\{\sin^2\left(\frac{k_{\min}\alpha}{2}\right),\epsilon_{\max}(k_{\min}\alpha,L_1(\alpha),L_2(\alpha))\right\}>0$ and the function $\epsilon_{\max}$ is defined in Lemma A.1. Then, only k satisfies (8).

Note that the conditions of Proposition 4.6 are satisfied by data obtained at time t_{crit} with $d'_{ij_1}(t_{\text{crit}})=0$, as in Lemma 4.5. The question now is to determine in what interval around t_{crit} the measured inter-drogue distance and distance derivative data still satisfies the conditions of Proposition 4.6. Among other things, this issue is important in order to determine acceptable sampling rates for the drogues. The next result answers this question.

Corollary 4.7 (Range of suitable times for determining k): Assuming internal wave parameters are within the bounds in Assumption 3.1, consider noiseless inter-drogue distance and distance derivative measurements at t_{crit} such that $d'_{ij_1}(t_{\text{crit}})=0$ and initial conditions satisfying

Assumption 4.3. Then, k uniquely satisfies (8) with data $\mathbf{D}(t)$, for all $t \in (t_{\text{crit}}-\Delta, t_{\text{crit}}+\Delta)$, where $\Delta(\delta, L_3, L_4) = \frac{L_4}{L_3}\frac{\delta}{1+\delta}$, δ is given in Proposition 4.6, $L_3 \geq 2\frac{\omega a}{k z_{\text{u}}}(\frac{\omega a}{z_{\text{u}}}+\omega)$, and $0 < L_4 \leq |d'_{ij_2}(t_{\text{crit}})|$.

Proof: The magnitude of the second time derivative of any inter-drogue distance is bounded by $2\frac{\omega a}{k z_{\text{u}}}(\frac{\omega a}{z_{\text{u}}} + \omega)$. Thus,

$$|d'_{ij_1}(t) - d'_{ij_1}(t_{\text{crit}})| \leq L_3|t - t_{\text{crit}}|, \qquad |d'_{ij_2}(t) - d'_{ij_2}(t_{\text{crit}})| \leq L_3|t - t_{\text{crit}}|.$$

From the analysis in Proposition 4.6, for a set of inter-drogue measurements there exists an open interval $(-\delta, \delta)$ in $\frac{d'_{ij_1}}{d'_{ij_2}}$ containing 0 where drq is strictly increasing. Thus, by the assumption that $t \in (t_{\text{crit}} - \Delta, t_{\text{crit}} + \Delta)$, we have the following,

$$\left|\frac{d'_{ij_1}(t + t_{\text{crit}})}{d'_{ij_2}(t + t_{\text{crit}})}\right| \leq \frac{L_3|t - t_{\text{crit}}|}{L_4 - L_3|t - t_{\text{crit}}|} < \frac{\frac{L_4\delta}{1+\delta}}{L_4(1 - \frac{\delta}{1+\delta})} = \delta,$$

which proves the result. ■

Using Corollary 4.7 and the assumptions in Assumptions 3.1 and 4.3, the following result gives a sufficient sampling rate to satisfy the conditions of Corollary 4.7.

Lemma 4.8 (Minimum sampling rate): If internal wave parameters satisfy Assumption 3.1 and given $\alpha > 0$ from Assumption 4.3, a bound on the minimum sampling rate for Corollary 4.7 is

$$f_{s,\min} > \frac{(1 + (\frac{a}{z_{\text{u}}})_{\max})\omega_{\max}}{\sin^2(\frac{k_{\min}\alpha}{2})\min\{\sin^2(\frac{k_{\min}\alpha}{2}), \epsilon_{\max}(\frac{k_{\min}\alpha}{2}, L_1(\alpha), L_2(\alpha))\}}. \tag{9}$$

C. Amplitude ratio and frequency via data fitting

In this section, we discuss how once the true horizontal wavenumber k is known, the parameters $\frac{a}{z_{\text{u}}}$ and ω can also be found as described in the following result.

Lemma 4.9 (Determination of $\frac{a}{z_u}$ and ω): Assume k is known. For $t_{\xi_1} < t_{\xi_2} < t_{\xi_3}$ with $t_{\xi_3} - t_{\xi_1} < T$, compute noiseless measurements of ψ_i and ψ'_i at these times by evaluating (5) and using the method described in Appendix A. Then, ω and $\frac{a}{z_{\text{u}}}$ can be found from

$$\begin{bmatrix}\beta_1 \\ \beta_2\end{bmatrix} = \begin{bmatrix}\sin(\psi_i(t_\varpi)) & 1 \\ \sin(\psi_i(t_\aleph)) & 1\end{bmatrix}^{-1}\begin{bmatrix}\psi'_i(t_\varpi) \\ \psi'_i(t_\aleph)\end{bmatrix} \quad \omega = -\beta_2, \quad \frac{a}{z_{\text{u}}} = \frac{-\beta_1}{\beta_2}, \tag{10}$$

where $\varpi, \aleph \in \{\xi_1, \xi_2, \xi_3\}$ such that $\sin(\psi_i(t_\varpi)) \neq \sin(\psi_i(t_\aleph))$.

Proof: First, given that $\psi'_i < 0$ and $\psi_i(t + T) = \psi_i(t) - 2\pi$, $\sin(\psi_i(t_\varpi)) \neq \sin(\psi_i(t_\aleph))$ for some $t_\varpi, t_\aleph \in \{t_{\xi_1}, t_{\xi_2}, t_{\xi_3}\}$. Using the values of ψ_i and ψ'_i at these two timesteps, one can solve for $\frac{a}{z_{\text{u}}}$ and ω using (4b) as described in the statement. ■

Remark 4.10 (Minimum sampling rate): For Lemma 4.9 to hold, one needs $f_{s,\min} > \frac{3}{T}$. Comparing this to Lemma 4.8, one can see that if (9) is enforced, then the assumptions of both Corollary 4.7 and Lemma 4.9 hold. •

D. Vanishing Distance Derivative Detection Strategy

We gather the discussion above into Algorithm 1.

Algorithm 1: Vanishing Distance Derivative Detection Strategy

Assumption: $f_s \geq f_{s,\min}$, initial distances satisfy Assumption 4.3 for some $\alpha > 0$, internal wave parameters within bounds in Assumption 3.1

run at time $t_\kappa = \frac{\kappa}{f_s}$, for some $\kappa \in \mathbb{Z}_{\geq 1}$

1 calculate wave propagation direction, $\theta_i = \text{atan}(-d_{ij_1}^{y_i\prime}(t_\kappa)/d_{ij_1}^{x_i\prime}(t_\kappa))$

2 **if** $\left|\frac{d'_{ij_1}(t_\kappa)}{d'_{ij_2}(t_\kappa)}\right| < \min\left\{\sin^2\left(\frac{k_{\min}\alpha}{2}\right), \epsilon_{\max}(k_{\min}\alpha, L_1(\alpha), L_2(\alpha))\right\}$ **then**

3 find k by solving $\text{drq}(\mathcal{k}, \mathbf{D}(t_\kappa)) = 0$

4 compute $\psi_i(t_\xi)$ (via (5)) and $\psi_i'(t_\xi)$ (via Appendix A), for $\xi \in \{\kappa-2, \kappa-1, \kappa\}$

5 choose $\varpi, \aleph \in \{\kappa-2, \kappa-1, \kappa\}$ such that $\sin(\psi_i(\varpi)) \neq \sin(\psi_i(\aleph))$ and solve

$$\begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix} = \begin{bmatrix}\sin(\psi_i(\varpi)) & 1\\ \sin(\psi_i(\aleph)) & 1\end{bmatrix}^{-1}\begin{bmatrix}\psi_i'(\varpi)\\ \psi_i'(\aleph)\end{bmatrix}$$

6 set $\omega = -\beta_2$ and $\frac{a}{z_u} = \frac{-\beta_1}{\beta_2}$

7 **end**

Note that for Step **3**, any root finder method can be used to find k uniquely; one suitable method, for instance, is gradient descent. The following result establishes the correctness of this strategy, which follows from Corollary 4.7, Lemma 4.9, and Remark 4.10.

Proposition 4.11 (Conditions for determining all parameters): Assuming that $f_s \geq f_{s,\min}$, internal wave parameters are within bounds in Assumption 3.1, and that the initial drogue locations satisfy Assumption 4.3, then drogue i can determine the parameters θ_i, $\frac{a}{z_u}$, ω, and k uniquely by using the Vanishing Distance Derivative Detection Strategy.

V. Robustness of parameter estimation under error

Here, we consider the effect of error in measurements on the application of the `Vanishing Distance Derivative Detection Strategy`. Section V-A describes some of the sources of error which occur during an ocean implementation of the proposed algorithm. Section V-B shows that the `Vanishing Distance Derivative Detection Strategy` is able to get a parameter estimates when the data has sufficiently small errors. This motivates our results in Section V-C, which bound the errors in estimates of k, $\frac{a}{z_u}$, and ω as a function of the errors in the measured quantities. Finally, in Section V-D, we devise a method for aggregating noisy parameter estimates from different timesteps.

A. Sources of error from algorithm implementation

Here, we describe some of the sources of error which occur in the algorithm's implementation.

Noise in measurements: In practice one can expect noise in measurements collected from sensors. We assume the sensor noise is unbiased, Gaussian, and that noise at different time instances and for different measurements are uncorrelated.

Model uncertainty: The problem setup described in Section III-B assumes that drogues are Lagrangian. In practice, as seen in Remark 3.2, drogues have a finite mass and drag coefficient making them not perfectly Lagrangian, leading to a difference between the actual drogue's velocity and the ocean velocity. One can treat this mismatch as an unknown but nonrandom error in the measurements of inter-drogue distances and distance derivatives.

Drogues not maintaining depth: We assume that the drogues have a controller that uses feedback on depth measurements to maintain a desired depth. Due to noisy depth measurements and a desire to minimize actuation cost, instead we assume that the drogues will be within an interval around the desired depth. Although depth is not directly used by the proposed algorithm, this inaccuracy affects inter-drogue distance measurements. As above, one can treat this as an unknown but nonrandom error in the inter-drogue distance measurements.

B. Existence of parameter estimates for measurements with error

Here we show that the `Vanishing Distance Derivative Detection Strategy` is able to estimate the parameters from measurements with sufficiently small error. We begin

our study with the horizontal wavenumber k because an estimate of it is needed for estimates of the other parameters. The next result establishes the analytic character of the function drq. The proof follows from the known fact, see e.g., [22], that sums, products, and compositions of analytic functions are analytic, and quotients of analytic functions are analytic, provided the denominator does not vanish.

Lemma 5.1 (drq *is analytic):* For any $k \in [k_{\min}, k_{\max}]$, drq is analytic on the set $\mathcal{D}_{\text{altc}}(k)$:

$$\mathcal{D}_{\text{altc}}(k) = \{\mathbf{D} \mid d'_{ij_1} \sin^2(\frac{kd_{ij_2}}{2}) - d'_{ij_2} \sin^2(\frac{kd_{ij_1}}{2}) \neq 0, \sin(\frac{kd_{ij_4}}{2}) \neq 0, \\ \cos(\frac{kd_{ij_4}}{2} + \psi_i(k, d_{ij_1}, d_{ij_2}, d'_{ij_1}, d'_{ij_2})) \neq 0, d'_{ij_4} \neq 0\}.$$

We now introduce two sets which help define the set of distances and distance derivatives $\mathbf{D}$ where estimates of k can be found in a neighborhood around $\mathbf{D}$. Let

$$\mathcal{D}_{\text{real}}(\mathbf{\Phi}) = \Big\{\mathbf{D} \mid 0 < d_{ij_1} < d_{ij_2} < d_{ij_3} < d_{ij_4} < \frac{2\pi}{k}, \\ d'_{ij} = 2\frac{\omega a}{k z_{\text{u}}} \sin\Big(\frac{kd_{ij}}{2}\Big) \cos\Big(\frac{kd_{ij}}{2} + \psi_i(k, d_{ij_1}, d_{ij_2}, d'_{ij_1}, d'_{ij_2})\Big),\ j \in \{j_1, \ldots, j_4\}\Big\}.$$

be the set of all inter-drogue measurements $\mathbf{D}$ that can come from one instantiation of $\mathbf{\Phi} = (\frac{a}{z_{\text{u}}}, \omega, k)$. Let $\mathcal{D}_{\text{diff}}(k) = \{\mathbf{D} \mid \partial_k \mathrm{drq}(k, \mathbf{D}) \neq 0\}$. Combining Lemma 5.1 with the Analytic Implicit Function Theorem [22] yields the existence of the implicit function for estimates of k.

Lemma 5.2 (Existence of estimates of horizontal wavenumber): For any $\mathbf{D} \in \mathcal{D}_{\text{real}}(\mathbf{\Phi}) \cap \mathcal{D}_{\text{diff}}(k) \cap \mathcal{D}_{\text{altc}}(k)$, there is a neighborhood of $\mathbf{D}$, $\mathcal{N}_{\mathbf{D}}(\mathbf{\Phi}) \subset \mathbb{R}^8$ for which there exists an analytic function $\mathbf{k_D} : \mathbb{R}^8 \to \mathbb{R}$ which satisfies

$$\mathrm{drq}(\mathbf{k_D}(\tilde{\mathbf{D}}), \tilde{\mathbf{D}}) = 0, \text{for } \tilde{\mathbf{D}} \in \mathcal{N}_{\mathbf{D}}(\mathbf{\Phi}).$$

The existence of the function $\mathbf{k_D}$ guarantees that k can be estimated from inter-drogue measurements containing errors, when the errors are sufficiently small.

Remark 5.3: (Frequency and amplitude ratio estimates from measurements with errors) From (5), given an estimate of k and measurements with errors, one can get an estimate of ψ_i. Furthermore, using the method outlined in Appendix A, one can also estimate ψ'_i. Thus, using (10), estimates for ω and $\frac{a}{z_{\text{u}}}$ exist as long as $\sin(\tilde{\psi}_i(t_\varpi)) \neq \sin(\tilde{\psi}_i(t_\aleph))$. •

C. Robustness to error

In this section we bound the error in estimates of the horizontal wavenumber k, amplitude ratio $\frac{a}{z_{\text{u}}}$, and frequency ω as a function of the error in the measurements.

1) Horizontal wavenumber: From the analysis in Section V-B, for a fixed set of noiseless measurements $\mathbf{D} \in \mathcal{D}_{\text{real}}(\mathbf{\Phi}) \cap \mathcal{D}_{\text{diff}}(k) \cap \mathcal{D}_{\text{altc}}(k)$, the corresponding noisy estimate of k in a neighborhood around $\mathbf{D}$, $\mathcal{N}_{\mathbf{D}}(\mathbf{\Phi})$, is given by the function $\mathbf{k_D}$. We wish to now restrict ourselves to a set where changes in the function $\mathbf{k_D}$ are bounded. Specifically, let $U_{\text{bnd-drv}} > 0$ and define

$$\mathcal{D}_{\text{deriv},U_{\text{bnd-drv}}}(\mathbf{\Phi}) = \{\mathbf{D} \in \mathcal{D}_{\text{real}}(\mathbf{\Phi}) \cap \mathcal{D}_{\text{diff}}(k) \cap \mathcal{D}_{\text{altc}}(k) \mid \max_{r\in\{1,\dots,8\}} |\partial_{\text{cpnt}_r(\mathbf{D})}\mathbf{k_D}(\mathbf{D})| < U_{\text{bnd-drv}}\}.$$

For each $\mathbf{D} \in \mathcal{D}_{\text{deriv},U_{\text{bnd-drv}}}(\mathbf{\Phi})$, by the analyticity of $\mathbf{k_D}$, one can construct a neighborhood $\mathcal{N}_{\mathbf{D},U_{\text{bnd-drv}}}(\mathbf{\Phi})$ such that for $\max_{r\in\{1,\dots,8\}} |\partial_{\text{cpnt}_r(\mathbf{D})}\mathbf{k_D}(\tilde{\mathbf{D}})| < U_{\text{bnd-drv}}$ for any $\tilde{\mathbf{D}} \in \mathcal{N}_{\mathbf{D},U_{\text{bnd-drv}}}(\mathbf{\Phi})$. The set of measurements with small enough error are $\tilde{\mathcal{D}}_{U_{\text{bnd-drv}}}(\mathbf{\Phi}) = \bigcup_{\mathbf{D}\in\mathcal{D}_{\text{deriv},U_{\text{bnd-drv}}}} \mathcal{N}_{\mathbf{D},U_{\text{bnd-drv}}}(\mathbf{\Phi})$.

The next result provides bounds on the error in estimating k by the error in the measurements. The proof is a combination of the Mean Value Theorem and the Cauchy-Schwartz inequality [23].

Lemma 5.4 (Bounds for errors in k as function of errors in measurements): Given noisy measurements of inter-drogue distances and distance derivatives $\tilde{\mathbf{D}} \in \tilde{\mathcal{D}}_{U_{\text{bnd-drv}}}(\mathbf{\Phi})$ for some $U_{\text{bnd-drv}} > 0$, then the error between the estimated $\hat{k}$ produced by the `Vanishing Distance Derivative Detection Strategy` and k can be bounded by

$$|\hat{k} - k| = |\mathbf{k_D}(\tilde{\mathbf{D}}) - k| \leq \sqrt{8}U_{\text{bnd-drv}}||\tilde{\mathbf{D}} - \mathbf{D}||.$$

2) Amplitude ratio and frequency: As seen in Lemma 4.9, with noiseless measurements, two measurements of ψ_i and ψ_i' are sufficient to exactly determine $\frac{a}{z_{\text{u}}}$ and ω. However, with noisy measurements, the question that naturally arises is whether there is a benefit to using more than 2 measurements. This is what we explore next.

Given n noisy measurements of ψ_i',

$$\tilde{\psi}_i'(t_{\kappa_q}) = \psi_i'(t_{\kappa_q}) + \epsilon_{\psi_i'}(t_{\kappa_q}), \quad \forall q \in \{1,\dots,n\},$$

we will construct estimates of $\frac{a}{z_{\text{u}}}$ and ω using least-squares techniques on the ψ_i dynamics in (4b),

$$\underbrace{\begin{bmatrix}\psi_i'(t_{\kappa_1})\\ \vdots \\ \psi_i'(t_{\kappa_n})\end{bmatrix}}_{\psi_i'} = \underbrace{\begin{bmatrix}\sin(\psi_i(t_{\kappa_1})) & 1\\ \vdots & \\ \sin(\psi_i(t_{\kappa_n})) & 1\end{bmatrix}}_{W} \underbrace{\begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix}}_{\beta}, \quad \beta_1 = \frac{\omega a}{z_{\text{u}}}, \quad \beta_2 = -\omega,$$

mimicking the technique in (10). The least-squares estimates are

$$\hat{\beta} = (W^T W)^{-1} W^T \tilde{\psi}_i' = (W^T W)^{-1} W^T (\psi_i' + \epsilon_{\psi_i'}) = \beta_{\text{true}} + \underbrace{(W^T W)^{-1} W^T \epsilon_{\psi_i'}}_{\beta_{\text{error}}}.$$

Explicitly, β_{error} is given by

$$\text{cpnt}_1(\beta_{\text{error}}) = \frac{\sum_{q=1}^n \epsilon_q (\sum_{r=1}^n \sin(\psi_i(t_{\kappa_q})) - \sin(\psi_i(t_{\kappa_r})))}{\sum_{q=1}^n \sin(\psi_i(t_{\kappa_q}))(\sum_{r=1}^n \sin(\psi_i(t_{\kappa_q})) - \sin(\psi_i(t_{\kappa_r})))},$$
$$\text{cpnt}_2(\beta_{\text{error}}) = \frac{\sum_{q=1}^n \epsilon_q (\sum_{r=1}^n \sin(\psi_i(t_{\kappa_r}))(\sin(\psi_i(t_{\kappa_r})) - \sin(\psi_i(t_{\kappa_q}))))}{\sum_{q=1}^n \sin(\psi_i(t_{\kappa_q}))(\sum_{r=1}^n \sin(\psi_i(t_{\kappa_q})) - \sin(\psi_i(t_{\kappa_r})))}.$$

The error β_{error} is a complex function of the distribution of ψ_i' as well as the sampling pattern (both spacing and the number of samples). Because the distribution of ψ_i' is non-Gaussian and unknown, we consider the case that all errors in ψ_i' are at most ϵ. Furthermore, we assume that the sampling pattern is uniform, meaning that

$$\sin(\psi_i(t_{\kappa_q})) = \frac{q-1}{n-1}(\sin(\psi_i(t_{\kappa_n})) - \sin(\psi_i(t_{\kappa_1}))) + \sin(\psi_i(t_{\kappa_1})), \quad \forall q \in \{2, n-1\}. \tag{11}$$

Let WLS be the *Worst-case Least Squares* error (for estimating ω), defined by,

$$\text{WLS}(n, \sin(\psi_i(t_{\kappa_1})), \sin(\psi_i(t_{\kappa_n})), \epsilon) = \epsilon \frac{\sum_{q=1}^n |(\sum_{r=1}^n \sin(\psi_i(t_{\kappa_r}))(\sin(\psi_i(t_{\kappa_r})) - \sin(\psi_i(t_{\kappa_q}))))|}{\sum_{q=1}^n \sin(\psi_i(t_{\kappa_q}))(\sum_{r=1}^n \sin(\psi_i(t_{\kappa_q})) - \sin(\psi_i(t_{\kappa_r})))}.$$

In the right-hand side, we use (11) to express $\{\sin(\psi_i(t_{\kappa_q}))\}_{q=2}^{n-1}$ in terms of $\sin(\psi_i(t_{\kappa_1}))$ and $\sin(\psi_i(t_{\kappa_n}))$. Note that WLS is an upper bound of $\text{cpnt}_2(\beta_{\text{error}})$. Even though the asymptotic dependence of WLS on n is difficult to characterize, the next result provides bounds that are sufficient to answer the question that motivates this section. Its proof is given in the appendix.

Lemma 5.5 (Worst-case estimation error grows with number of measurements): Consider any maximum error $\epsilon > 0$, wave parameters $\frac{a}{z_u}, \omega \in \mathbb{R}$, number of measurements $n \in \mathbb{Z}_{\geq 1}$, and range of measurements $\sin(\psi_i(t_{\kappa_1})) < \sin(\psi_i(t_{\kappa_n})) \in [-1, 1)$. If the set of measurements of ψ_i are distributed according to (11) and the errors in ψ_i' are bounded by ϵ, i.e.,

$$|\tilde{\psi}_i'(t_{\kappa_q}) - \psi_i'(t_{\kappa_q})| < \epsilon, \quad \forall q \in \{1, \ldots, n\},$$

then WLS can be bounded between two increasing functions of n as

$$\epsilon \frac{3 \max\{|\sin(\psi_i(t_{\kappa_1}))|, |\sin(\psi_i(t_{\kappa_n}))|\}(n-1)}{(\sin(\psi_i(t_{\kappa_n})) - \sin(\psi_i(t_{\kappa_1})))n} \geq \text{WLS}(n, \sin(\psi_i(t_{\kappa_1})), \sin(\psi_i(t_{\kappa_n})), \epsilon)$$
$$\geq \begin{cases} \epsilon & -1 < \sin(\psi_i(t_{\kappa_1})) \leq 0, \\ \epsilon \max\left(\frac{3 \sin(\psi_i(t_{\kappa_1}))(n+3)(n-2)}{2(\sin(\psi_i(t_{\kappa_n})) - \sin(\psi_i(t_{\kappa_1})))n(n+1)}, 1\right) & 0 < \sin(\psi_i(t_{\kappa_1})) < 1. \end{cases}$$

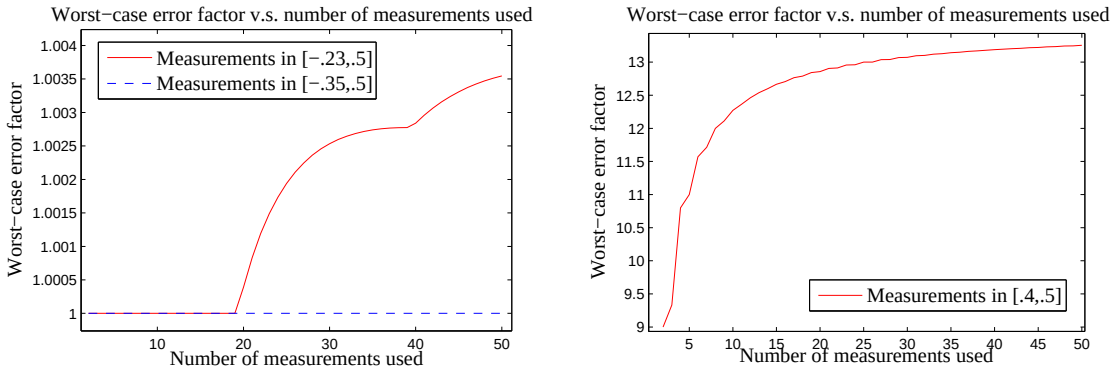


Fig. 5. In the uniformly distributed measurements case, the worst case error factor (amount multiplying ϵ) is an increasing function of n.

Lemma 5.5 is validated by Figure 5 which shows that WLS is an increasing function of n under uniformly distributed measurements. Given that the drogues' measurements are roughly uniformly distributed, that the distribution of errors in ψ_i' is non-Gaussian and not necessarily unbiased, and Lemma 5.5, we use two as the number of measurements to estimate ω and $\frac{a}{z_u}$.

D. Parameter estimate aggregation

In this section we consider the task of aggregating parameter estimates. Before we can tackle this, we must define the distributions for measurements. Here, we assume that the the non-random errors stemming from the dynamical assumptions are small relative to the random sensor noise. As seen in Remark 3.2, the mismatch in dynamical models rapidly becomes small. The error associated with the depth mismatch can be made small by choosing a tighter depth-keeping interval. Thus, given an inter-drogue distance d and distance derivative d', we let $\tilde{d}$, $\tilde{d}'$, denote the measurements of d and d' by a drogue, with the following Gaussian error model

$$\tilde{d} = d + \epsilon_d, \quad \epsilon_d \sim \mathcal{N}(0, \sigma_d^2), \tag{12a}$$

$$\tilde{d}' = d' + \epsilon_{d'}, \quad \epsilon_{d'} \sim \mathcal{N}(0, \sigma_{d'}^2). \tag{12b}$$

We assume that the variances are a function of the specific sensors used and are known.

Within one period of the wave, the `Vanishing Distance Derivative Detection Strategy` generates many estimates of the parameters (an estimate is generated every time that the condition in Step **2** of Algorithm 1 is satisfied). Furthermore, drogues may be sampling over

the course of several periods. Therefore, it makes sense to improve the estimation by fusing together estimates obtained at different timesteps. However, the synthesis of the appropriate fusion mechanism is challenging because the distribution of estimates of the parameters is non-Gaussian due to the nonlinearity of the dynamics and the operations within the algorithm. This is the problem that we tackle next, beginning with an informal description.

> *[Informal description]:* Because the parameters' distributions are only implicitly defined and non-Gaussian, we create an approximation up to a desired order $p \in \mathbb{Z}_{\geq 1}$. Using on the fact that measurements are Gaussian, we compute the expectation and variance of this approximate distribution and them to properly fuse parameter estimates.

Given the fact that $\mathbf{k_D}$ is analytic on $\mathcal{N}_{\mathbf{D}}(\mathbf{\Phi})$, we use Taylor series to generate approximations of arbitrary order. Formally, given an arbitrary analytic function $\mathbf{pr_D} : \mathbb{R}^8 \to \mathbb{R}$ with $\mathbf{pr_D}(\mathbf{D}) = \mathrm{pr}$, the pth-order Taylor series expression around $\mathbf{D}$ is given by

$$\mathbf{pr_D}(\tilde{\mathbf{D}}) = \mathrm{pr} + T^p_{\mathbf{D}}(\tilde{\mathbf{D}}) + R^p_{\mathbf{D}}(\tilde{\mathbf{D}}), \tag{13}$$

where

$$T^p_{\mathbf{D}}(\tilde{\mathbf{D}}) = \sum_{q=1}^{p} \frac{1}{q!} \mathbf{pr}^{(q)}_{\mathbf{D}}(\mathbf{D}; \tilde{\mathbf{D}} - \mathbf{D}), \quad R^p_{\mathbf{D}}(\tilde{\mathbf{D}}) = \mathbf{pr}^{(p+1)}_{\mathbf{D}}(\mathbf{D}^*; \tilde{\mathbf{D}} - \mathbf{D}),$$

for some $\mathbf{D}^* \in [\tilde{\mathbf{D}}, \mathbf{D}]$ and parameter value pr, where

$$\mathbf{pr}^{(q)}_{\mathbf{D}}(\mathbf{D}; \tilde{\mathbf{D}} - \mathbf{D}) = \sum_{r_1=1}^{8} \cdots \sum_{r_q=1}^{8} \frac{\partial^q \mathbf{pr_D}(\tilde{\mathbf{D}})}{\partial_{\mathrm{cpnt}_{r_1}(\mathbf{D})} \cdots \partial_{\mathrm{cpnt}_{r_q}(\mathbf{D})}} \cdot$$
$$(\mathrm{cpnt}_{r_1}(\tilde{\mathbf{D}}) - \mathrm{cpnt}_{r_1}(\mathbf{D})) \ldots (\mathrm{cpnt}_{r_q}(\tilde{\mathbf{D}}) - \mathrm{cpnt}_{r_q}(\mathbf{D})).$$

For $\mathbf{pr_D} = \mathbf{k_D}$, equation (13) represents the noisy parameter calculated from a set of noisy measurements $\tilde{\mathbf{D}}$. Since Taylor's Theorem provides existence of the truncation term but no constructive way to determine it, we seek to investigate how accurate the pth order approximation $T^p_{\mathbf{D}}(\tilde{\mathbf{D}})$ is. Specifically, the form of the function, along with knowing the Gaussian distribution of the measurements, allows one to calculate the expectation and variance of this (approximate) distribution of the parameter. In practice, one cannot quite calculate these quantities, because they require partial derivatives which must be evaluated at the noiseless measurements $\mathbf{D}$, which are not available. However, by using $\tilde{\mathbf{D}}$ one may approximately determine these quantities.

Determining the expectation and variance of individual parameter estimates with the method described above allows us to devise a strategy to fuse them to get a more accurate approximation. Formally, given independent random variables x_1 and x_2 with mean $\mathrm{E}[x_1] = \mathrm{E}[x_2] = \mu$ and variances $\mathrm{Var}[x_1] = \sigma_1^2$, $\mathrm{Var}[x_2] = \sigma_2^2$, consider the optimal aggregating function OptAgg by

$$\mathrm{OptAgg}(x_1, \sigma_1^2, x_2, \sigma_2^2) = \left(\frac{\frac{x_1}{\sigma_1^2} + \frac{x_2}{\sigma_2^2}}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}}, \frac{1}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}} \right).$$

Here, $\mathrm{cpnt}_1(\mathrm{OptAgg})$ is the new random variable and $\mathrm{cpnt}_2(\mathrm{OptAgg})$ is its variance. This is the convex combination of x_1 and x_2 that results in the random variable with the smallest variance.

We are now ready to define the p`th-Order Parameter Fusion` procedure. Given a sequence of noisy parameter estimates $\{(\hat{\mathrm{pr}}_\ell, \tilde{\mathbf{D}}_\ell) \mid \ell \in \mathbb{Z}_{\geq 1}\}$ determined from noisy measurements, $\hat{\mathrm{pr}}_\ell = \mathbf{pr}_{\mathbf{D}_\ell}(\tilde{\mathbf{D}}_\ell)$, this procedure generates a sequence of estimates $\{\widehat{\mathrm{pr}}^p_\ell \mid \ell \in \mathbb{Z}_{\geq 1}\}$ by means of the following iterative aggregation process

$$(\widehat{\mathrm{pr}}^p_{\ell+1}, \mathrm{Var}[\widehat{\mathrm{pr}}^p_{\ell+1}]) = \mathrm{OptAgg}(\widehat{\mathrm{pr}}^p_\ell, \mathrm{Var}[\widehat{\mathrm{pr}}^p_\ell], \hat{\mathrm{pr}}_{\ell+1} - \mathrm{E}[T^p_{\mathbf{D}_{\ell+1}}(\tilde{\mathbf{D}}_{\ell+1})], \mathrm{Var}[T^p_{\mathbf{D}_{\ell+1}}(\tilde{\mathbf{D}}_{\ell+1})]), \quad (14)$$

where $\widehat{\mathrm{pr}}^p_1 = \hat{\mathrm{pr}}_1 - \mathrm{E}[T^p_{\mathbf{D}_1}(\tilde{\mathbf{D}}_1)]$ and $\mathrm{Var}[\widehat{\mathrm{pr}}^p_1] = \mathrm{Var}[T^p_{\mathbf{D}_1}(\tilde{\mathbf{D}}_1)]$. According to this procedure, the pth-order estimate $\widehat{\mathrm{pr}}^p$ is sequentially updated by optimally combining the previous aggregated value with the next parameter estimate (after the expected bias has been removed). The next result, whose proof is given in the appendix, establishes its convergence under suitable conditions on the pth-order approximation of $\mathbf{pr}$.

Proposition 5.6 (pth-order aggregation): For noisy inter-drogue measurements $\{\tilde{\mathbf{D}}_\ell \mid \ell \in \mathbb{Z}_{\geq 1}\}$ containing additive Gaussian noise according to (12), assume there exist $\epsilon_E \geq 0$ and $\epsilon_V \geq 0$ such that the following bounds hold uniformly for all $\ell \in \mathbb{Z}_{\geq 1}$,

$$|\,\mathrm{E}[\mathbf{pr}_{\mathbf{D}_\ell}(\tilde{\mathbf{D}}_\ell) - T^p_{\mathbf{D}_\ell}(\tilde{\mathbf{D}}_\ell)] - \mathrm{pr}\,| \leq \epsilon_E, \quad \mathrm{Var}[\mathbf{pr}_{\mathbf{D}_\ell}(\tilde{\mathbf{D}}_\ell)] \leq \epsilon_V.$$

Then, the iterates (14) generated by p`th-Order Parameter Fusion` satisfy

$$\lim_{\ell \to \infty} \Pr[|\widehat{\mathrm{pr}}^p_\ell - \mathrm{pr}\,| \leq \epsilon_E + \epsilon] = 1, \quad \forall \epsilon > 0.$$

Note that for $p = 1$, one is estimating the distribution of the parameter as a sum of Gaussian distributions because there are only first-order terms in the function T^1 in (13). Similarly for $p = 2$, the distribution is the sum of Gaussian distributions plus second-order Chi-squared distributions. Chi-squared distributions have non-zero expectation, and so, the `Second-Order`

k `Fusion` outperforms the `First-Order` k `Fusion` because it estimates what the distribution's bias is and subtracts this from individual parameter estimates before fusing them.

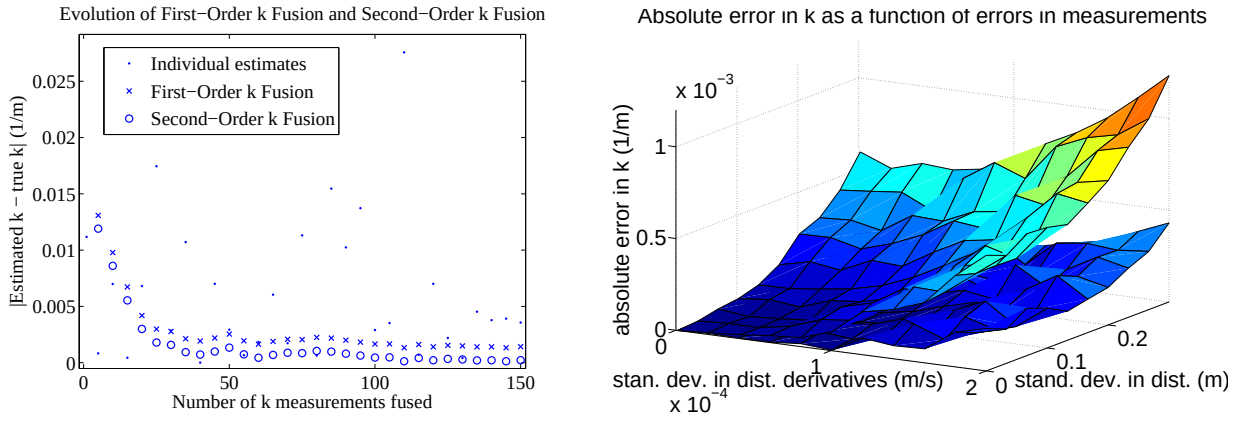
Figure 6(a) shows the evolution of p`th-Order Parameter Fusion` for $p = 1$ and $p = 2$ for estimates of k generated by the `Vanishing Distance Derivative Detection Strategy`. The x-axis corresponds to the number of estimates of k fused together. Each k estimate is obtained at a different instant of time. Note that both evolutions, after only a small number of fusions, have a smaller error than the individual measurements, and that the evolution corresponding to $p = 2$ has a smaller asymptotic error. Note that Proposition 5.6 is not directly applicable to make guarantees on convergence because the implicit functions that give estimates of k (cf. Lemma 5.2) and ω and $\frac{a}{z_{\mathrm{u}}}$ (cf. Remark 5.3) have domains that are not all $\mathbb{R}^8$. This implies that the Gaussian noise may occasionally be too large to produce estimates. However, as the standard deviation of the measurement noise get smaller, the fraction of acceptable noisy measurements increases and so the execution of the p`th-Order Parameter Fusion` more closely mirrors Proposition 5.6. One can see that the simulations are in line with the result.

Figure 6(b) shows the absolute error of p`th-Order Parameter Fusion` for $p = 1$ and $p = 2$, using k estimates from the `Vanishing Distance Derivative Detection Strategy`, as a function of the standard deviation in inter-drogue distance and distance derivatives measurements, which depicts that $p = 2$ outperforms $p = 1$.

Finally, Figure 7 compares an inter-drogue trajectory generated from true wave parameters with the trajectory that would have occurred from estimated parameters. Specifically, the `Vanishing Distance Derivative Detection Strategy` and the `Second-Order` k `Fusion` method are used to generate and fuse estimates of k after which the other parameters are estimated. One can see that the trajectory closely tracks the true trajectory.

VI. Conclusions

This paper has considered the task of estimating the physical parameters of a horizontally propagating ocean linear internal wave using a group of drogues. We have established an explicit analytic description of the evolution of a drogue under the flow induced by the linear internal wave. This result implies that inter-drogue distances evolve in a purely periodic way. We have built on this knowledge to design the `Vanishing Distance Derivative Detection Strategy`. This strategy relies on the fact that inter-drogue distance derivatives become close



(a) (b)

Fig. 6. (a) shows pth-Order Parameter Fusion applied to estimates of k from Vanishing Distance Derivative Detection Strategy, for $p = 1$ and $p = 2$. After only a few estimates are fused, the error is already much smaller than the individual estimates. Also, $p = 2$ converges to a smaller error than $p = 1$. (b) shows the absolute error of pth-Order Parameter Fusion applied to estimates of k from Vanishing Distance Derivative Detection Strategy as a function of the standard deviations in measurement noise, highlighting how $p = 2$ outperforms $p = 1$. In both figures the parameter values, taken from [1], are $k = \frac{2\pi}{190}\frac{1}{m}$, $\frac{a}{z_u} = \frac{1}{7}$, and $\omega = \frac{2\pi}{960}\frac{1}{s}$. The drogues are in a line, initially $10m$ apart from the closest drogues on either side. In (a), the standard deviations in distances and distance derivatives are $.01m$ and $.0005\frac{m}{s}$, respectively. In (b) for each set of standard deviations, $10,000$ estimates were fused.

to zero multiple times during one period. Under noiseless measurements, we have established that the algorithm exactly computes the internal wave parameters and derived conditions on the minimal sampling rate for this to happen. Next, we have characterized the robustness of our strategy. Under measurements with error, we have bounded the error in the parameter estimates as a function of the errors in the measured quantities. For the case of measurements corrupted by additive Gaussian noise, we have also developed a general scheme termed pth-Order Parameter Fusion for aggregating parameter estimates based on determining the pth-order approximation of their distribution. The method results in smaller errors than the individual estimates generated by the Vanishing Distance Derivative Detection Strategy.

Future work will be devoted to the extension of our algorithm in several directions: to include information from the control actuation employed to maintain the depth of the drogue, to consider general drogue dynamics not necessarily Lagrangian, and to study scenarios with multiple internal waves present, including the possibility of weakly nonlinear waves. We are currently exploring the practical implementation of our approach in a network of drogues under development at the

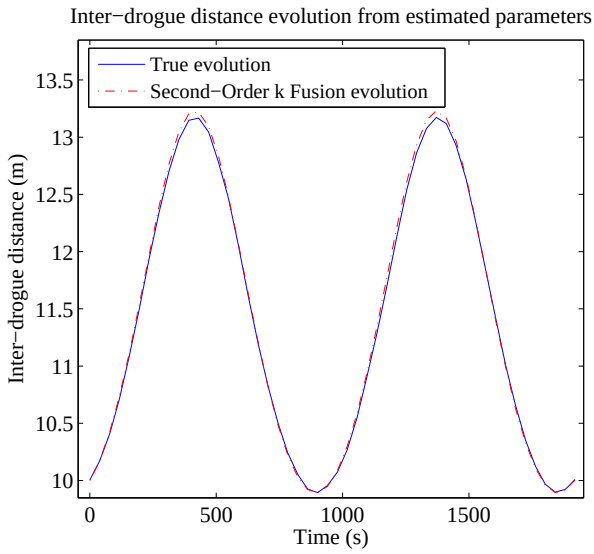


Fig. 7. Comparison between inter-drogue trajectories generated from a set of true wave parameters and from the parameters estimated by the `Vanishing Distance Derivative Detection Strategy` and the `Second-Order` k `Fusion` method. Here, the standard deviations in distances and distance derivatives are $.01m$ and $.0001\frac{m}{s}$, respectively. The estimated trajectory closely matches the true trajectory.

UCSD Scripps Institute of Oceanography [24]. Finally, we wish to determine general conditions that allow us to employ the proposed algorithm in a broader class of dynamical systems.

REFERENCES

[1] C. Lennert-Cody and P. Franks, "Plankton patchiness in high-frequency internal waves," *Marine Ecology Progress Series*, vol. 18, pp. 59–66, 1999.

[2] J. R. Zeldis and J. B. Jillett, "Aggregation of pelagic Mundia gregana (Fabncius) (Decapoda, Anomura) by coastal fronts and internal waves," *Journal of Plankton Research*, vol. 4, no. 4, pp. 839–857, 1982.

[3] A. L. Shanks, "Surface slicks associated with tidally forced internal waves may transport pelagic larvae of benthic invertebrates and fishes shoreward," *Marine Ecology Progress Series*, vol. 13, pp. 311–315, 1983.

[4] P. Franks, "Spatial patterns in dense algal blooms," *Limnology and Oceanography*, vol. 42, no. 5, pp. 1297–1305, 1997.

[5] R. Susanto, L. Mitnik, and Q. Zheng, "Ocean internal waves observed in the Lombok Strait," *Oceanography*, vol. 18, no. 4, pp. 80–87, 2005.

[6] M. Perry and D. Rudnick, "Observing the ocean with autonomous and Lagrangian platforms and sensors," *Oceanography*, vol. 16, no. 4, pp. 31–36, 2003.

[7] H. J. Freeland and P. F. Cummins, "Argo: A new tool for environmental monitoring and assessment of the world's oceans, and example from the N.E. Pacific," *Progress in Oceanography*, vol. 64, no. 1, pp. 31–44, 2005.

[8] Y. Han, R. A. de Callafon, J. Cortés, and J. Jaffe, "Dynamic modeling and pneumatic switching control of a submersible drogue," in *International Conference on Informatics in Control, Automation and Robotics*, vol. 2, (Funchal, Madeira, Portugal), pp. 89–97, June 2010.

[9] J. Jouffroy, Q. Zhou, and O. Zielinski, "Towards selective tidal-stream transport for Lagrangian profilers," in *Oceans*, (Waikoloa, HI), Sept. 2011.

[10] N. E. Leonard, D. Paley, F. Lekien, R. Sepulchre, D. M. Fratantoni, and R. Davis, "Collective motion, sensor networks, and ocean sampling," *Proceedings of the IEEE*, vol. 45, no. 1, pp. 48–74, 2007. Special Issue on Networked Control Systems.

[11] D. Paley, F. Zhang, and N. Leonard, "Cooperative control for ocean sampling: the glider coordinated control system," *IEEE Transactions on Control Systems Technology*, vol. 16, no. 4, pp. 735–744, 2008.

[12] R. Graham and J. Cortés, "Adaptive information collection by robotic sensor networks for spatial estimation," *IEEE Transactions on Automatic Control*, vol. 57, no. 6, pp. 1404–1419, 2012.

[13] Y. Ru and S. Martínez, "Coverage control in constant flow environments based on a mixed energy-time metric," *Automatica*, 2012. Submitted.

[14] L. DeVries and D. Paley, "Multi-vehicle control in a strong flowfield with application to hurricane sampling," *AIAA Journal of Guidance, Control, and Dynamics*, vol. 35, no. 3, pp. 794–806, 2012.

[15] G. Seber and C. Wild, *Nonlinear Regression*. Wiley, 1989.

[16] K. Schittkowski, *Data Fitting in Dynamical Systems*. Kluwer Academic Publishers, 2002.

[17] M. Craymer, *The least squares spectrum, its inverse transform and autocorrelation function: theory and some applications in Geodesy*. PhD thesis, Univ. of Toronto, 1998.

[18] W. Emery and R. Thomson, *Data Analysis Methods in Physical Oceanography*. Elsevier, 2001.

[19] A. Gill, *Atmosphere-Ocean Dynamics*. Academic Press, 1982.

[20] C. Woolsey, "Vehicle dynamics in currents," Technical Report 2011-01, Virginia Polytechnic Institute & State University, Blacksburg,VA, Sept. 2011.

[21] C. Brennen, "A review of added mass and fluid inertial forces," tech. rep., Naval Civil Engineering Laboratory, 1982.

[22] S. G. Krantz and H. R. Parks, *A Primer of Real Analytic Functions*. Birkhäuser Advanced Texts, Boston, MA: Birkhäuser, 2nd ed., 2002.

[23] W. Rudin, *Principles of Mathematical Analysis*. McGraw-Hill, 1953.

[24] Jaffe Laboratory for Underwater Imaging. Scripps Institute of Oceanography, "Autonomous Underwater Explorer project." La Jolla, CA. http://jaffeweb.ucsd.edu.

[25] A. Savitzky and M. Golay, "Smoothing and differentiation of data by simplified least squares procedures," *Analytical Chemistry*, vol. 36, no. 8, pp. 1627–1639, 1964.

APPENDIX

The appendix contains some basic results used in the paper and the proofs of the main results.

A. Derivative estimation from noisy data

Here, we consider estimating an analytic function $f : \mathbb{R} \to \mathbb{R}$ and its first (time) derivative $f' : \mathbb{R} \to \mathbb{R}$ from n evenly sampled measurements in the sampling window T, when the measurements

are corrupted by additive Gaussian noise. So that our method is causal, at any given time we only use the n most recent measurements. Additionally, for computational reasons, at every timestep, we relabel the current timestep as $t = 0$ and the times of all other measurements accordingly, i.e., we have measurements for times $\{t_\zeta = \frac{-\zeta}{n-1}T\}_{\zeta\in\{0,\dots,n-1\}}$. The noisy measurement at t_ζ is

$$\tilde{f}(t_\zeta) = f(t_\zeta) + \epsilon(t_\zeta), \quad \epsilon \sim \mathcal{N}(0, \sigma^2).$$

The method we use is a polynomial smoothing filter approach [25] because this allows us to justify that the derivative estimates are unbiased Gaussian random variables. For $p \ll n \in \mathbb{Z}_{\geq 1}$, we construct a pth-order polynomial filter from n evenly spaced noisy measurements over the sampling window T using the data $\{(t_\zeta, \tilde{f}(t_\zeta))\}_{\zeta\in\{0,\dots,n-1\}}$. Consider the Taylor series expansion,

$$\underbrace{\begin{bmatrix} \tilde{f}(t_0) \\ \tilde{f}(t_1) \\ \vdots \\ \tilde{f}(t_{n-1}) \end{bmatrix}}_{\tilde{\mathbf{F}}} = \underbrace{\begin{bmatrix} 1 & 0 & \dots & 0 \\ 1 & t_1 & \dots & t_1^p \\ \vdots & \vdots & \vdots & \vdots \\ 1 & t_{n-1} & \dots & t_{n-1}^p \end{bmatrix}}_{V} \underbrace{\begin{bmatrix} f(t_0) \\ f'(t_0) \\ \vdots \\ \frac{1}{p!}f^{(p)}(t_0) \end{bmatrix}}_{\mathbf{G}} + \underbrace{\begin{bmatrix} 0 \\ \sum_{j=p+1}^{\infty} f^{(j)}(t_0)\frac{t_1^j}{j!} \\ \vdots \\ \sum_{j=p+1}^{\infty} f^{(j)}(t_0)\frac{t_{n-1}^j}{j!} \end{bmatrix}}_{\epsilon_{\text{bias}}} + \underbrace{\begin{bmatrix} \epsilon_0 \\ \epsilon_1 \\ \vdots \\ \epsilon_{n-1} \end{bmatrix}}_{\epsilon_{\text{random}}}.$$

More compactly, this can be written as $\tilde{\mathbf{F}} = V\mathbf{G} + \epsilon_{\text{bias}} + \epsilon_{\text{random}}$. The least-squares estimate for $f(t_0)$ and $f'(t_0)$ are the first, $\text{cpnt}_1(\mathbf{G})$, and second, $\text{cpnt}_2(\mathbf{G})$, components of the vector $\mathbf{G}$,

$$\hat{f}(t_0) = \text{cpnt}_1(\mathbf{G}) = \text{cpnt}_1((V^TV)^{-1}V^T(\tilde{\mathbf{F}} - \epsilon_{\text{bias}} - \epsilon_{\text{random}})),$$
$$\hat{f}'(t_0) = \text{cpnt}_2(\mathbf{G}) = \text{cpnt}_2((V^TV)^{-1}V^T(\tilde{\mathbf{F}} - \epsilon_{\text{bias}} - \epsilon_{\text{random}})).$$

We ignore the bias which arises from considering only the pth-order expansion of f because, for a fixed p, it can be made arbitrarily close to zero by choosing the sampling window T small enough. With this observation in mind, the estimate of $f'(t_0)$ is an (unbiased) Gaussian random variable with variance $\sigma_{f'}^2 = \text{cpnt}_{2,2}((V^TV)^{-1})\sigma^2$.

B. Proofs of results from Sections IV and V

Proof of Proposition 4.1: Let $\psi(t) = kp^x(t) - \omega t + \phi$ be the relative phase between the wave and drogue. Then (2) can be written as $\psi' = \omega\left(\frac{a}{z_u}\sin\psi - 1\right)$. Integrating, one gets

$$\int_{\psi_0}^{\psi} \frac{d\varsigma}{\frac{a}{z_u}\sin\varsigma - 1} = \omega\int_0^t d\tau,$$
$$\frac{2}{\sqrt{1-(a/z_u)^2}}\text{atan}\left(\frac{a/z_u - \tan(\zeta/2)}{\sqrt{1-(a/z_u)^2}}\right)\Big|_{\psi_0}^{\psi} = \omega t,$$

where $\psi_0 = \psi(0) = kp^x(0) + \phi$. Manipulating the last expression, we arrive at

$$\psi(t) = 2\text{atan}\left(\frac{a}{z_{\text{u}}} - \sqrt{1 - \left(\frac{a}{z_{\text{u}}}\right)^2}\tan\left(\sqrt{1-(a/z_{\text{u}})^2}\frac{\omega t}{2} + \Lambda_0\right)\right), \tag{15}$$

$$\Lambda_0 = \text{atan}\left(\frac{1}{\sqrt{1-(a/z_{\text{u}})^2}}\left(\frac{a}{z_{\text{u}}} - \tan\left(\frac{kp^x(0)+\phi}{2}\right)\right)\right).$$

Note, however, that this function is discontinuous or, in other words, the expression for $\psi(t)$ above is only valid if the argument of the tangent function is in the interval $[-\frac{\pi}{2}, \frac{\pi}{2}]$. A general expression for $\psi(t)$ can be obtained as follows. Since the period of the tangent function is π, we deduce that the fundamental period T of ψ is $T = \frac{2\pi}{\omega\sqrt{1-(a/z_{\text{u}})^2}}$. Note that at

$$t = \frac{n\pi - 2\Lambda_0}{\omega\sqrt{1-(a/z_{\text{u}})^2}}, \tag{16}$$

with $n \in \mathbb{Z}$ odd, one has that $\frac{\sqrt{1-(a/z_{\text{u}})^2}}{2}\omega t + \Lambda_0 = \frac{n\pi}{2}$, and hence (15) jumps from $-\pi$ to π. So, to obtain an expression of $\psi(t)$ which is valid in general, we need to subtract 2π from (15) every time t crosses one of the critical times (16) or, in other words, subtract the quantity

$$2\pi\left\lfloor \frac{t - \frac{1}{\omega\sqrt{1-(a/z_{\text{u}})^2}}(\pi - 2\Lambda_0)}{T} + 1\right\rfloor = 2\pi\left\lfloor \frac{t}{T} + \frac{\Lambda_0}{\pi} + \frac{1}{2}\right\rfloor.$$

Finally, note that the initial condition Λ_0 jumps from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$ at $kp^x(0) + \phi = n\pi$ for $n \in \mathbb{Z}$ odd. Thus, in order to make Λ_0 change continuously with the initial conditions, we subtract away from Λ_0 the quantity $\pi\left\lfloor \frac{kp^x(0)+\phi+\pi}{2\pi}\right\rfloor$. The result now follows. ■

The following auxiliary result is needed before we present the proof of Proposition 4.6. We begin by defining $\epsilon_{\text{max}} : \mathbb{R}^3_{>0} \to \mathbb{R}_{>0}$,

$$\epsilon_{\text{max}}(x, C_1, C_2) = \max_{\gamma\in[0,\Gamma(x)]} \mathcal{R}(x, \gamma, C_1, C_2) > 0, \tag{17}$$

with $\Gamma(x) = x - \sin(x)\cos(x)$ and

$$\mathcal{R}(x, \gamma, C_1, C_2) = \begin{cases} \min\left\{\frac{\frac{1}{2}\arcsin(2(x-\gamma))-x}{C_1}, \frac{\gamma}{C_2}\right\} & \frac{\pi}{4} > x - \gamma > 0, \\ \frac{x-\gamma-\frac{1}{2}}{C_2} & 2\pi > x - \gamma > \frac{\pi}{4}. \end{cases}$$

Lemma A.1: For any $x \in (0, 2\pi)$, $C_1, C_2 \in \mathbb{R}_{>0}$, $\epsilon \in (-\epsilon_{\text{max}}(x, C_1, C_2), \epsilon_{\text{max}}(x, C_1, C_2))$,

$$\frac{1}{2}\sin(2(x + C_1\epsilon)) - x + C_2\epsilon \leq 0, \tag{18}$$

Proof: Note that for any $x > 0$, $\gamma \in (0, \Gamma(x))$ and $C_1, C_2 \in \mathbb{R}_{>0}$, $R(x, \gamma, C_1, C_2) > 0$, ensuring that $\epsilon_{\max}(x, C_1, C_2) > 0$ as well. For γ such that $x - \gamma > \frac{\pi}{4}$, we know that for any $\epsilon \in [-\mathcal{R}(x, \gamma, C_1, C_2), \mathcal{R}(x, \gamma, C_1, C_2)]$,

$$\frac{1}{2}\sin(2(x + C_1\epsilon)) - x + C_2\epsilon \leq \frac{1}{2} - x + \frac{C_2(x - \gamma - \frac{1}{2})}{C_2} = -\gamma \leq 0.$$

For γ such that $0 < x - \gamma \leq \frac{\pi}{4}$, note that for any $\epsilon \in [-\mathcal{R}(x, \gamma, C_1, C_2), \mathcal{R}(x, \gamma, C_1, C_2)]$,

$$\frac{1}{2}\sin(2(x + C_1\epsilon)) - x + C_2\epsilon \leq \frac{1}{2}\sin(2(x + C_1\frac{\frac{1}{2}\arcsin(2(x-\gamma)) - x}{C_1})) - x + \gamma = 0,$$

which completes the result. ■

Proof of Proposition 4.6: The proof proceeds by establishing $\partial_{\mathcal{k}} \operatorname{drq}(\mathcal{k}, \mathbf{D}) > 0$, for all $\mathcal{k} \in [k_{\min}, k_{\max}]$. Once this is shown, it is easy to see that only k satisfies (8) since drq is strictly increasing as a function of $\mathcal{k}$. To prove this fact about drq, it is enough to establish that

$$\frac{\partial_{\mathcal{k}} \operatorname{dr}(\mathcal{k}, \boldsymbol{\nu}(\mathcal{k}, d_{ij_1}, d_{ij_2}, d'_{ij_1}, d'_{ij_2}), d_{ij_3})}{\operatorname{dr}(\mathcal{k}, \boldsymbol{\nu}(\mathcal{k}, d_{ij_1}, d_{ij_2}, d'_{ij_1}, d'_{ij_2}), d_{ij_3})} > \frac{\partial_{\mathcal{k}} \operatorname{dr}(\mathcal{k}, \boldsymbol{\nu}(\mathcal{k}, d_{ij_1}, d_{ij_2}, d'_{ij_1}, d'_{ij_2}), d_{ij_4})}{\operatorname{dr}(\mathcal{k}, \boldsymbol{\nu}(\mathcal{k}, d_{ij_1}, d_{ij_2}, d'_{ij_1}, d'_{ij_2}), d_{ij_4})},$$

as long as

$$\operatorname{sgn}(\operatorname{dr}(\mathcal{k}, \boldsymbol{\nu}(\mathcal{k}, d_{ij_1}, d_{ij_2}, d'_{ij_1}, d'_{ij_2}), d_{ij_3})) = \operatorname{sgn}(\operatorname{dr}(\mathcal{k}, \boldsymbol{\nu}(\mathcal{k}, d_{ij_1}, d_{ij_2}, d'_{ij_1}, d'_{ij_2}), d_{ij_4}), \tag{19}$$

which corresponds to

$$\boldsymbol{\nu} \in \Big(\frac{-\pi}{2} - \frac{\mathcal{k} d_{ij_3}}{2}, \frac{\pi}{2} - \frac{\mathcal{k} d_{ij_4}}{2}\Big) \bigcup \Big(\frac{\pi}{2} - \frac{\mathcal{k} d_{ij_3}}{2}, \frac{3\pi}{2} - \frac{\mathcal{k} d_{ij_4}}{2}\Big).$$

This set admits all $\frac{d'_{ij_1}}{d'_{ij_2}} \in \Big(-\infty, \frac{\sin(\frac{\mathcal{k} d_{ij_1}}{2})\sin(\frac{\mathcal{k}}{2}(d_{ij_4} - d_{ij_1}))}{\sin(\frac{\mathcal{k} d_{ij_2}}{2})\sin(\frac{\mathcal{k}}{2}(d_{ij_4} - d_{ij_2}))}\Big)$. By hypothesis, $\frac{d'_{ij_1}}{d'_{ij_2}}$ is in this range and hence (19) holds. Since $d_{ij_3} < d_{ij_4}$, its sufficient to show that

$$\partial_{d_{ij_3}}\left(\frac{\partial_{\mathcal{k}} \operatorname{dr}(\mathcal{k}, \boldsymbol{\nu}(\mathcal{k}, d_{ij_1}, d_{ij_2}, d'_{ij_1}, d'_{ij_2}), d_{ij_3})}{\operatorname{dr}(\mathcal{k}, \boldsymbol{\nu}(\mathcal{k}, d_{ij_1}, d_{ij_2}, d'_{ij_1}, d'_{ij_2}), d_{ij_3})}\right) < 0.$$

After some calculations, we obtain

$$\partial_{d_{ij_3}}\left(\frac{\partial_{\mathcal{k}} \operatorname{dr}(\mathcal{k}, \boldsymbol{\nu}(\mathcal{k}, d_{ij_1}, d_{ij_2}, d'_{ij_1}, d'_{ij_2}), d_{ij_3})}{\operatorname{dr}(\mathcal{k}, \boldsymbol{\nu}(\mathcal{k}, d_{ij_1}, d_{ij_2}, d'_{ij_1}, d'_{ij_2}), d_{ij_3})}\right) = \tag{20}$$

$$\frac{\sin\left(\frac{\mathcal{k} d_{ij_3}}{2}\right)\cos\left(\frac{\mathcal{k} d_{ij_3}}{2}\right) - \frac{\mathcal{k} d_{ij_3}}{2}}{2\sin^2\left(\frac{\mathcal{k} d_{ij_3}}{2}\right)} - \frac{\sin\left(\frac{\mathcal{k} d_{ij_3}}{2} + \boldsymbol{\nu}\right)\cos\left(\frac{\mathcal{k} d_{ij_3}}{2} + \boldsymbol{\nu}\right) + \mathcal{k}\left(\frac{d_{ij_3}}{2} + \partial_{\mathcal{k}}\boldsymbol{\nu}\right)}{2\cos^2\left(\frac{\mathcal{k} d_{ij_3}}{2} + \boldsymbol{\nu}\right)}.$$

From Lemma 4.5, when $\frac{d'_{ij_1}}{d'_{ij_2}} = 0$, then $\nu = \frac{\pi}{2} - \frac{\kappa d_{ij_1}}{2}$ and $\partial_\kappa \nu = -\frac{d_{ij_1}}{2}$, and so (20) becomes

$$\partial_{d_{ij_3}}\left(\frac{\partial_\kappa \operatorname{dr}(\kappa,\nu(\kappa,d_{ij_1},d_{ij_2},d'_{ij_1},d'_{ij_2}),d_{ij_3})}{\operatorname{dr}(\kappa,\nu(\kappa,d_{ij_1},d_{ij_2},d'_{ij_1},d'_{ij_2}),d_{ij_3})}\right) = \frac{\sin\left(\frac{\kappa d_{ij_3}}{2}\right)\cos\left(\frac{\kappa d_{ij_3}}{2}\right) - \frac{\kappa d_{ij_3}}{2}}{2\sin^2\left(\frac{\kappa d_{ij_3}}{2}\right)} + \frac{\sin\left(\frac{\kappa(d_{ij_3}-d_{ij_1})}{2}\right)\cos\left(\frac{\kappa(d_{ij_3}-d_{ij_1})}{2}\right) - \kappa\frac{(d_{ij_3}-d_{ij_1})}{2}}{2\sin^2\left(\frac{\kappa(d_{ij_3}-d_{ij_1})}{2}\right)},$$

where both summands of which are clearly negative ensuring the desired partial derivative is negative. With this in mind, we write in general that $\nu = \frac{\pi}{2} - \frac{\kappa d_{ij_1}}{2} + \epsilon_1$ and $\partial_\kappa \nu = -\frac{d_{ij_1}}{2} + \epsilon_2$. Similarly, we rewrite (20) as

$$\partial_{d_{ij_3}}\left(\frac{\partial_\kappa \operatorname{dr}(\kappa,\nu(\kappa,d_{ij_1},d_{ij_2},d'_{ij_1},d'_{ij_2}),d_{ij_3})}{\operatorname{dr}(\kappa,\nu(\kappa,d_{ij_1},d_{ij_2},d'_{ij_1},d'_{ij_2}),d_{ij_3})}\right) = \frac{\sin\left(\frac{\kappa d_{ij_3}}{2}\right)\cos\left(\frac{\kappa d_{ij_3}}{2}\right) - \frac{\kappa d_{ij_3}}{2}}{2\sin^2\left(\frac{\kappa d_{ij_3}}{2}\right)} + \frac{\sin\left(\frac{\kappa(d_{ij_3}-d_{ij_1})}{2}+\epsilon_1\right)\cos\left(\frac{\kappa(d_{ij_3}-d_{ij_1})}{2}+\epsilon_1\right) - \kappa\frac{(d_{ij_3}-d_{ij_1})}{2} - \kappa\epsilon_2}{2\sin^2\left(\frac{\kappa(d_{ij_3}-d_{ij_1})}{2}+\epsilon_1\right)}.$$

The unchanged, first summand is still negative. A sufficient condition for the whole expression being negative is that the second summand is negative too. Using Lemma A.1, we can ensure that (20) is negative when $\left|\frac{d'_{ij_1}}{d'_{ij_2}}\right| < \epsilon_{\max}\left(\frac{\kappa(d_{ij_3}-d_{ij_1})}{2}, L_1, L_2\right)$, which completes the expression for δ. The fact that $\delta > 0$ is seen from the definition of $\epsilon_{\max}$ and the assumption that $\kappa \le \frac{2\pi}{d_{ij_4}}$. ■

Proof of Lemma 5.5: To reduce the length of expressions in the proof, we rewrite the data with the following notation,

$$\tilde{y}_q = \beta_1 x_q + \beta_2 + \epsilon_q, \qquad x_q = \frac{q-1}{n-1}(x_n - x_1) + x_1, \quad \forall q \in \{1,\dots,n\},$$

$$x_1 = \sin(\psi_i(t_{\kappa_1})), \quad x_n = \sin(\psi_i(t_n)), \quad x_q = \sin(\psi_i(t_{\kappa_q})),$$
$$\tilde{y}_q = \tilde{\psi}'_i(t_{\kappa_q}), \quad \epsilon_q = \tilde{\psi}'_i(t_{\kappa_q}) - \psi'_i(t_{\kappa_q}), \quad \beta_1 = \frac{\omega a}{z_{\mathrm{u}}}, \quad \beta_2 = -\omega.$$

Then,

$$\begin{aligned}\operatorname{WLS}(n,x_1,x_n,\epsilon) &= \epsilon\frac{\sum_{q=1}^n \left|\frac{x_n-x_1}{n-1}\sum_{r=1}^n(\frac{r-1}{n-1}(x_n-x_1)+x_1)(q-r)\right|}{\sum_{q=1}^n\left((\frac{q-1}{n-1}(x_n-x_1)+x_1)(\frac{x_n-x_1}{n-1}\sum_{r=1}^n(q-r))\right)} \\ &\le \frac{\epsilon\frac{x_n-x_1}{n-1}\max\{|x_1|,|x_n|\}\sum_{q=1}^n\sum_{r=1}^n\left|q-r\right|}{\frac{1}{12}\frac{(x_n-x_1)^2}{n-1}n^2(n+1)} \\ &= \epsilon\frac{3\max\{|x_1|,|x_n|\}(n-1)}{(x_n-x_1)n},\end{aligned}$$

which is clearly strictly increasing in n because $\sum_{q=1}^{n}\sum_{r=1}^{n}|r-q| = \frac{n(n+1)(n-1)}{3}$. WLS can be bounded below by

$$\mathrm{WLS}(n,x_1,x_n,\epsilon) \geq \frac{\epsilon|\sum_{q=1}^{n}(\sum_{r=1}^{n}x_r(x_r-x_q))|}{\sum_{q=1}^{n}x_q(\sum_{r=1}^{n}x_q-x_r)} = \epsilon$$

for any $x_1 < x_n \in [-1,1)$. Additionally, for any $0 < x_1 < x_n < 1$,

$$\begin{aligned}
\mathrm{WLS}(n,x_1,x_n,\epsilon) &= \epsilon\frac{\frac{x_n-x_1}{n-1}}{\sum_{q=1}^{n}\left((\frac{q-1}{n-1}(x_n-x_1)+x_1)(\frac{x_n-x_1}{n-1}\sum_{r=1}^{n}(q-r)\right)}\cdot \\
&\quad \sum_{q=1}^{n}\Big|\sum_{r=1}^{q-1}(\frac{r-1}{n-1}(x_n-x_1)+x_1)(q-r)+\sum_{r=q+1}^{n}(\frac{r-1}{n-1}(x_n-x_1)+x_1)(q-r)\Big| \\
&\geq \epsilon\frac{\frac{x_n-x_1}{n-1}\sum_{q=1}^{\frac{n}{2}-1}\left|\sum_{r=2q}^{n}x_1(q-r)\right|}{\sum_{q=1}^{n}\left((\frac{q-1}{n-1}(x_n-x_1)+x_1)(\frac{x_n-x_1}{n-1}\sum_{r=1}^{n}(q-r)\right)} \\
&= \epsilon\frac{|x_1|\frac{x_n-x_1}{n-1}\frac{n(n+3)(n-2)}{8}}{\frac{1}{12}\frac{(x_n-x_1)^2}{n-1}n^2(n+1)} = \epsilon\frac{3|x_1|(n+3)(n-2)}{2(x_n-x_1)n(n+1)},
\end{aligned}$$

which completes the result. ■

Proof of Proposition 5.6: First, we note that $\widehat{\mathrm{pr}}^p_{\ell+1}$ can be written in the following non-recursive way:

$$\widehat{\mathrm{pr}}^p_{\ell+1} = \frac{\sum_{q=1}^{\ell+1}\frac{\mathbf{pr}_{\mathbf{D}_q}(\tilde{\mathbf{D}}_q)-\mathrm{E}[T^p_{\mathbf{D}_q}(\tilde{\mathbf{D}}_q)]}{\mathrm{Var}[T^p_{\mathbf{D}_q}(\tilde{\mathbf{D}}_q)]}}{\sum_{q=1}^{\ell+1}\frac{1}{\mathrm{Var}[T^p_{\mathbf{D}_q}(\tilde{\mathbf{D}}_q)]}}.$$

Thus, the variance of $\widehat{\mathrm{pr}}^p_{\ell+1}$ is

$$\mathrm{Var}[\widehat{\mathrm{pr}}^p_{\ell+1}] = \frac{\sum_{q=1}^{\ell+1}\frac{\mathrm{Var}\left[\mathbf{pr}_{\mathbf{D}_q}(\tilde{\mathbf{D}}_q)-\mathrm{E}[T^p_{\mathbf{D}_q}(\tilde{\mathbf{D}}_q)]\right]}{\mathrm{Var}[T^p_{\mathbf{D}_q}(\tilde{\mathbf{D}}_q)]^2}}{(\sum_{q=1}^{\ell+1}\frac{1}{\mathrm{Var}[T^p_{\mathbf{D}_q}(\tilde{\mathbf{D}}_q)]})^2} = \frac{\sum_{q=1}^{\ell+1}\frac{\mathrm{Var}\left[\mathbf{pr}_{\mathbf{D}_q}(\tilde{\mathbf{D}}_q)\right]}{\mathrm{Var}[T^p_{\mathbf{D}_q}(\tilde{\mathbf{D}}_q)]^2}}{(\sum_{q=1}^{\ell+1}\frac{1}{\mathrm{Var}[T^p_{\mathbf{D}_q}(\tilde{\mathbf{D}}_q)]})^2} \leq \frac{\epsilon_V\sum_{q=1}^{\ell+1}\frac{1}{\mathrm{Var}[T^p_{\mathbf{D}_q}(\tilde{\mathbf{D}}_q)]^2}}{(\sum_{q=1}^{\ell+1}\frac{1}{\mathrm{Var}[T^p_{\mathbf{D}_q}(\tilde{\mathbf{D}}_q)]})^2}.$$

From there, we notice that $\lim_{\ell\to\infty}\mathrm{Var}[\widehat{\mathrm{pr}}^p_\ell] = 0$. Now we bound the expected value of $\widehat{\mathrm{pr}}^p - \mathrm{pr}$:

$$|\,\mathrm{E}[\widehat{\mathrm{pr}}^p_{\ell+1}-\mathrm{pr}]| = \left|\mathrm{E}\left[\sum_{q=1}^{\ell+1}\frac{\frac{\mathbf{pr}_{\mathbf{D}_\ell}(\tilde{\mathbf{D}}_q)-\mathrm{E}(T^p_{\mathbf{D}_q}(\tilde{\mathbf{D}}_q))}{\mathrm{Var}(T^p_{\mathbf{D}_q}(\tilde{\mathbf{D}}_q))}}{\sum_{q=1}^{\ell+1}\frac{1}{\mathrm{Var}(\mathbf{pr}_{\tilde{\mathbf{D}}_q}(\tilde{\mathbf{D}}_q))}}-\mathrm{pr}\right]\right| \leq \sum_{q=1}^{\ell+1}\frac{\frac{|\,\mathrm{E}[\mathbf{pr}_{\mathbf{D}_\ell}(\tilde{\mathbf{D}}_q)-\mathrm{E}(T^p_{\mathbf{D}_q}(\tilde{\mathbf{D}}_q))-\mathrm{pr}]|}{\mathrm{Var}[T^p_{\mathbf{D}_q}(\tilde{\mathbf{D}}_q)]}}{\sum_{q=1}^{\ell+1}\frac{1}{\mathrm{Var}[\mathbf{pr}_{\tilde{\mathbf{D}}_q}(\tilde{\mathbf{D}}_q)]}} \leq \epsilon_E.$$

This implies that for all $\epsilon > 0$, $\lim_{\ell\to\infty}\Pr[|\widehat{\mathrm{pr}}^p_\ell - \mathrm{E}[\widehat{\mathrm{pr}}^p_\ell]| < \epsilon] = 1$. Also,

$$\begin{aligned}
\Pr[|\widehat{\mathrm{pr}}^p_\ell - \mathrm{E}[\widehat{\mathrm{pr}}^p_\ell]| < \epsilon] &\leq \Pr[|\widehat{\mathrm{pr}}^p_\ell - \mathrm{E}[\widehat{\mathrm{pr}}^p_\ell| + |\,\mathrm{E}[\widehat{\mathrm{pr}}^p_\ell - \mathrm{pr}]|] < \epsilon+\epsilon_E], \\
&= \Pr[|\widehat{\mathrm{pr}}^p_\ell - \mathrm{E}[\widehat{\mathrm{pr}}^p_\ell| + |\,\mathrm{E}[\widehat{\mathrm{pr}}^p_\ell] - \mathrm{pr}\,|] < \epsilon+\epsilon_E], \\
&\leq \Pr[|\widehat{\mathrm{pr}}^p_\ell - \mathrm{pr}\,| < \epsilon+\epsilon_E],
\end{aligned}$$

which shows the convergence result. ■