Partitioning ocean flow field for underwater vehicle path planning

Mengxue Hou, Haoyan Zhai, Homin Zhou, Fumin Zhang

***Abstract*—In this paper, we present a flow field partition method that extracts the key features, which are the spatial and temporal variation of the flow field. The partition method is developed based on K-means algorithm. In the case where the temporal pattern of the flow field contains only a periodic tidal component, we propose an algorithm that partitions the flow field into static clusters of piece-wise constant flow, by performing K-means clustering over the time-averaged flow field. Then the method is extended to partitioning the flow field into clusters of uniform time-varying flow, by fitting the spatially averaged flow in each static partitioned region to a parametric flow model. Simulation results of partitioning both a simulated jet flow field, as well as the ocean surface flow data into time-invariant and time-varying uniform flow are presented to demonstrate that the proposed method can represent the true flow field with significantly less amount of data. Result of using Method of Evolving Junctions to plan the time-optimal path in the partitioned flow field is also presented to demonstrate that the proposed flow partitioning method can be applied to facilitate path planning, and can reduce the path planning computational cost.**

I. Introduction

Autonomous Underwater Vehicles (AUVs) are a class of submerged marine robots that are able to perform persistent missions in the ocean. Over the last few decades, AUVs have been widely applied to various applications, including ocean sampling [1], [2], surveillance and inspection [3], and many more.

Along with the advancements in computation power and sensor technology equipped in AUVs, there is a growing trend on deploying multiple AUVs to perform the adaptive environmental sampling and sensing tasks together [4]–[6]. With the field estimation information being shared among all agents, the mobile sensor network can exploit their mobility to perform autonomous missions such as gradient climbing and feature tracking in an uncertain environment [7]. However, the amount of information that can be shared among vehicles is limited due to constraints in communication capabilities. This demands information reduction before its transmission. For example, the reduced flow maps, not the actual flow fields, are often shared [8]. On the other hand, information reduction of the flow map can also lead to significantly less computational cost for map merging and path planning. Similar map reduction idea has been proposed on path planning of mobile robot, where a reduced-element map representation of a grid map is generated to save path planning computational cost [9]. However, to the best of our knowledge, there is no existing work on applying the reduced-element flow map representation for path planning of AUV, and evaluating path planning computational cost with respect to the flow map reduction.

The fundamental requirement of flow map reduction is that the reduced map should preserve all the key patterns of the original map. The spatial and temporal variation of the flow field are of the most critical patterns of the flow field that need to be shared by the mobile network in order to perform path planning for the agents. The spatial and temporal variation pattern of a flow field can be described with reduced elements by partitioning the flow field into regions of uniform flow speed. The spatial variation of the flow field can thus be represented by the pattern of partitioned regions, while the temporal variation patterns are described by the temporal variation of the partitioned uniform flow.

Since the partitioned flow field preserves the key patterns of the flow field, path planning can be performed given these extracted feature information. Method of Evolving Junctions (MEJ) is a general approach to solve optimal control problem with constraints [10], and is recently applied to solve the path planning problem of AUV traveling in flow fields [11]. Given a partitioned configuration space, MEJ solves for the optimal path by recasting the infinite dimensional path planning problem into a finite dimensional optimization method, through introducing a number of junction points that can slide on the boundary of partitioned regions. Then the globally optimal path can be found by optimizing position of the junction points using intermittent diffusion method [12]. Comparing to other path planning methods, the following features of MEJ make it particularly applicable to path planning of underwater vehicle. First, MEJ significantly reduces the computation time by transforming the infinite dimensional problem into lower dimensions. Besides, MEJ guarantees to generate the global optimal path with probability one. Moreover, MEJ is applicable to path planning with various forms of cost function. These merits of the algorithm make it promising for application in AUV path planning.

In this paper, we propose a K-means based method for dividing a spatio-temporally varying flow field into static physically similar regions. Similarity is defined as a weighted sum of the Eulerian distance and flow difference between every two points in the domain. In the case where the temporal variation of the field contains only periodic tidal component, we propose to partition the flow field into time-invariant segments of piece-

Mengxue Hou and Fumin Zhang are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Haoyan Zhai and Haomin Zhou are with the School of Mathematics, Georgia Institute of Technology, {mhou30, hzhai8, hz25, fumin}@gatech.edu

wise constant flow. In the case where the flow field contains more complex temporal variation, we divide the flow field into static segments containing uniform time-varying flow. Using simulation, we demonstrate that the partitioned flow field shares sufficient similarity with the original flow field by comparing the Lagrangian particles trajectory in the original and the partitioned flow field. Also we show that partitioning the flow field reduces the path planning computational cost, and at the same time reduces the amount of information to represent a complicated realistic ocean surface flow field.

Flow field partition and information reduction techniques have been studied under the goal of visualizing and understanding the physical dynamics of the flow field. A detailed review of the existing techniques can be found in [13]. In general, the partitioning problem can be concluded into a clustering problem. The goal is to partition a set of data into groups, elements of each group should be as similar as possible. Most of the work solves the problem in an Eulerian point of view, defining the difference of data points as the flow velocity difference and distance between data points. [14] applies a hierarchical top-down approach based on clustering and principal component analysis, while [15] proposes a bottom-up method that merges clusters with high similarity. [16] uses a Voronoi tessellation based algorithm for clustering and derives the mean of regions. For most of the existing literature following the Eulerian point of view, partitioning is performed on the static flow map at each time step. Therefore, the map partition changes over time, and the time-varying segments do not tend to yield an obvious physical meaning [17]. To the best of our knowledge, there is only one work on partitioning the flow field into time-invariant segments. [18] considers partitioning time-varying flow field into static segments, by modeling the direction of the flow field observations as a von Mises-Fisher distribution based HMM (Hidden Markov Model), then computing partitions of the flow field by clustering the limiting distribution of HMM in the domain using spectral clustering. However, by partitioning the limiting distribution of HMM, this paper considers long-term patterns, and ignores the short-term dynamics of the flow field. This might make the method inapplicable to ocean flow partitioning for supporting AUV path planning. Since time interval of AUV missions is several magnitudes less than the time scale of ocean flow evolution, the short-term ocean flow dynamics provide critical information for AUV planning and control, and thus cannot be neglected.

In recent years, methods of flow field partitioning have also been developed following the Lagrangian point of view, with the difference between data point defined as separation of streamlines [19], [20]. In the case of path planning for AUV with low thrust capabilities, AUV dynamics shares great similarity with the dynamics of passive tracer. Therefore, the separation of particle trajectories provides significant insight to AUV path planning [21]. However, in the case where AUV speed is comparable to or exceeds flow speed, AUV dynamics will differ from passive tracer dynamics, and the separation of streamlines may not provide enough information for AUV path planning.

The contribution of this work are two-fold. First, we propose a method that extracts the time-invariant spatial patterns and temporal patterns of flow field dynamics within certain partitioning time interval. By adjusting the partitioning time interval, the extracted flow patterns can represent either the short-term or long-term flow dynamics. Second, we demonstrate that the extracted spatial and temporal patterns of the flow field can facilitate AUV path planning, and reduce the path planning computational cost.

The rest of the paper is organized as follows. The problem formulation, and introduction on the background of this paper are presented in Section II. Section III describes the algorithm for partitioning the flow field into piece-wise constant regions, while Section IV outlines the algorithm for dividing the flow field into piece-wise time-varying regions. Simulation results of flow partition and path planning are presented in Section V, while Section VI provides the conclusion of the paper.

II. Problem Formulation

Consider a $2D$ spatial domain $\mathcal{D}$, $\mathbf{F} : \mathcal{D} \times [T_0, T_f] \to \mathbb{R}^2$ denotes the flow field in $\mathcal{D}$ over a time period $[T_0, T_f]$. We assume $\mathbf{F}$ is available over a uniform grid on a set of discrete time steps, $T = \{T_0, T_1, \ldots, T_f\}$. Given $\mathbf{F}$, we wish to derive a simplified flow field containing a finite number of flow regions of uniform flow speed. The simplified flow field should preserve the spatial and temporal variation features of the true flow field.

Denoting the number of partitioned regions as k, we represent each partitioned flow region as $R_\alpha, \alpha \in \{1, 2, \ldots, k\}$, and the partitioned flow field to be $\hat{\mathbf{F}}(\mathbf{x}, t)$. We reduce the spatial variation information of the flow field by partitioning the field into regions, and ignore the detailed spatial variability inside the region. This lead to the following assumption.

Assumption II.1. *Given any α, $\hat{\mathbf{F}}(\mathbf{x}, t)$ is uniform $\forall \mathbf{x} \in R_\alpha$.*

Given the above assumption, we drop the dependency of $\hat{\mathbf{F}}(\mathbf{x}, t)$ on $\mathbf{x}$, and denote the partitioned flow in the partitioned regions as $\mathcal{F} = \{\hat{\mathbf{F}}_1(t), \ldots \hat{\mathbf{F}}_k(t)\}$.

A. K-means clustering

Let $\mathbf{y} \in \mathbb{R}^n$ denote the observation data. Given the number of clusters to be k, K-means partitions the flow field into k sets $\mathcal{R} = \{R_1, R_2, \ldots, R_k\}$ by minimizing the following intra-cluster variance function,

$$\min_{\mathcal{R}} \sum_{\alpha=1}^{k} \sum_{\mathbf{y} \in R_\alpha} \text{dist}(\mathbf{y}, \mu_\alpha),$$

where μ_α is the mean of observation points in R_α, $\text{dist} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ denotes a specific distance function. Commonly used distance functions include Euclidean distance, Manhattan distance, Minkowski distance, and many more [22].

To perform K-means algorithm, the mean of the k clusters are first initialized randomly. Then the algorithm iteratively performs two steps. The assignment step assigns each $\mathbf{y}$ to the nearest mean μ_α. The update step adjusts μ_α to match the

mean of observation data points that are assigned to the α^{th} cluster. The algorithm terminates when the assignment of data points does not change over iterations.

B. Method of Evolving Junctions

Considering a vehicle with dynamics described as $\dot{\mathbf{x}} = \hat{\mathbf{F}}(\mathbf{x}) + \mathbf{u}(\mathbf{x}, t)$, where $\hat{\mathbf{F}}$ represents a flow field divided into convex flow partitions of piece-wise constant flow, and $\mathbf{u}$ denotes the vehicle's forward speed, we aim to handle the following problem:

$$\begin{aligned} \min_{\mathbf{u},T} & \int_{t_0}^{t_f} L(\mathbf{u})dt \\ s.t. \quad & \dot{\mathbf{x}} = \hat{\mathbf{F}}(\mathbf{x}) + \mathbf{u}, \\ & \mathbf{x}(t_0) = \mathbf{x}_0, \\ & \mathbf{x}(t_f) = \mathbf{x}_f, \\ & \max_{t\in[0,T]} \|\mathbf{u}\| \leq V, \end{aligned} \tag{1}$$

where $L(\mathbf{u})$ is a general convex cost function with respect to $\mathbf{u}$. Further, we define the intersection between the vehicle path and the flow region boundaries as junction points. Since the flow is uniform inside each region, the optimal path in each region is a straight line. Thus the optimal path can be parameterized using positions of junctions. By rewriting the cost function into an explicit function of the junction points, MEJ transforms an infinite dimensional path planning problem into a finite dimensional optimization problem with respect to the junction positions. This optimization problem in general is not convex, and thus is solved by the intermittent diffusion method to guarantee global optimality [12].

III. Partitioning flow field into uniform constant clusters

In this section, we describe an algorithm to partition the flow field into static flow field containing regions of piece-wise constant flow. We make the following assumption in this section.

Assumption III.1. *The partitioned flow speed $\hat{\mathbf{F}}_\alpha$ in R_α is a constant vector over the partition interval T.*

Remark III.1. *The evolution of oceanography processes, including the mesoscale eddies and currents such as Gulf Stream, can be at the scale of weeks to month [23]. Such temporal scale is usually larger than the scale of an AUV deployment. Therefore, by proper selection of the partition time interval, the difference between the true flow and the partitioned flow contains approximately only the periodic tidal components. Since the effect of periodic tidal components on vehicle trajectory is bounded, partitioning the flow field into piece-wise constant flow regions will not lead to significant deviation of vehicle trajectory.*

Let $\mathbf{y}(t) \in \mathbb{R}^4, \mathbf{y}(t) = [\mathbf{x}; \mathbf{F}(\mathbf{x}, t)]$ denote the vector observation at position $\mathbf{x}$ and time t. Due to the mismatch of unit of distance and velocity, we introduce a diagonal positive definite weighting matrix $\mathbf{Q}$ that scales elements in both $\mathbf{x}$ and $\mathbf{F}(\mathbf{x}, t)$ to $[-1, 1]$. Then the squared distance function is defined as

$$\text{dist}^2(\mathbf{y}, \mathbf{y}') = (\mathbf{y} - \mathbf{y}')^T \mathbf{Q} (\mathbf{y} - \mathbf{y}').$$

Given the defined distance function, the objective of the clustering analysis is to find a solution to the following optimization problem,

$$\min_{\mathcal{R},\mathcal{M}} J = \sum_{\alpha=1}^{k} \sum_{\mathbf{y}\in R_\alpha} \sum_{t\in T} \text{dist}^2(\mathbf{y}(t), \mu_\alpha), \tag{2}$$

where $\mathcal{R} = \{R_1, \ldots, R_k\}$, $\mathcal{M} = \{\mu_1, \ldots, \mu_k\}$.

Claim III.1. *For any partition, the cost function J defined in Eq. (2) is equivalent to the following expression*

$$J' = \sum_{\alpha=1}^{k} \sum_{\mathbf{y}\in R_\alpha} \left[(T_f - T_0)\text{dist}^2(\bar{\mathbf{y}}, \mu_\alpha) + \sum_{t\in T} \text{dist}^2(\mathbf{y}(t), \bar{\mathbf{y}}) \right],$$

where $\bar{\mathbf{y}} = [\mathbf{x}; \bar{\mathbf{F}}(\mathbf{x})]$, $\bar{\mathbf{F}}(\mathbf{x}) = \frac{1}{T_f - T_0} \sum_{t\in T} \mathbf{F}(\mathbf{x}, t)$ represents the time-averaged flow speed over time interval T.

Proof. For a random selection of observation data $\mathbf{y}(t)$ in one partitioned region R_α, define $d\mathbf{y}(t) = \mathbf{y}(t) - \bar{\mathbf{y}}$. Note that since $\sum_{t\in T} d\mathbf{y}(t) = \sum_{t\in T} \mathbf{y}(t) - (T_f - T_0)\bar{\mathbf{y}} = 0$, the following equality holds

$$\begin{aligned} & \sum_{t\in T} \text{dist}^2(\mathbf{y}(t), \mu_\alpha) \\ =& \sum_{t\in T} \Big[\bar{\mathbf{y}}^T Q \bar{\mathbf{y}} + 2d\mathbf{y}(t)^T Q \bar{\mathbf{y}} + d\mathbf{y}(t)^T Q d\mathbf{y}(t) \\ & - 2\mu_\alpha^T Q \bar{\mathbf{y}} - 2\mu_\alpha^T Q d\mathbf{y}(t) + \mu_\alpha^T Q \mu_\alpha \Big] \\ =& (T_f - T_0)\text{dist}^2(\bar{\mathbf{y}}, \mu_\alpha) + \sum_{t\in T} \text{dist}^2(\mathbf{y}(t), \bar{\mathbf{y}}). \end{aligned}$$

□

Remark III.2. *Since the second term in J' represents the temporal variation of flow speed on one grid point, it does not change with respect to different partitioning of the flow field. Therefore, the optimal solution of Eq. (2) equals the optimal solution of*

$$\min_{\mathcal{R},\mathcal{M}} \sum_{\alpha=1}^{k} \sum_{\mathbf{y}\in R_\alpha} \text{dist}^2(\bar{\mathbf{y}}, \mu_\alpha). \tag{3}$$

Therefore, it can be concluded that the optimal partition of the time-varying flow field is equivalent to the optimal partition of the time-averaged flow field.

Implementing K-means algorithm requires knowing the number of partitioned regions prior to segmenting the field. If the field is divided into too many regions, then the algorithm actually fails on the intention of information reduction. On the other hand, dividing the field into too few regions may result in a large difference between the true and the partitioned flow field, which will lead to significant error in path planning results. Therefore, we introduce an iterative K-means algorithm

that can guarantee a bounded flow field partition error, and at the same time utilize the least number of partition regions.

Define the flow field partition error as

$$\delta F = \max_{\alpha \in 1,\dots,k} \max_{\mathbf{x} \in R_\alpha} ||\bar{\mathbf{F}}(\mathbf{x}) - \hat{\mathbf{F}}_\alpha||. \tag{4}$$

Given an initialized k, we will iteratively perform K-means, and check if the flow partition error satisfies $\delta F < \epsilon$. If the condition is satisfied, then the current k is designated for clustering. Otherwise the number of partition is increased by one. The algorithm is described in Algorithm 1.

Algorithm 1: Dividing spatio-temporally varying flow field into piece-wise constant flow field

Data: flow field observations $\mathbf{y}(t)$, flow partition time interval T, partitioning error threshold ϵ.
Output: flow field partitions $\mathcal{R}$, piece-wise constant flow speed of each region $\mathcal{F}$.

```
Compute time-averaged flow field ȳ over the time interval T;
Initialize cluster number k = 1, randomly initialize cluster centroid position;
Compute solution R, M to Eq. (3) using K-means method;
Compute the flow partition error δF using Eq. (4) ;
while δF > ε do
    k = k + 1;
    Randomly initialize cluster centroids;
    Compute solution R, M to Eq. (3) using K-means method;
end
```

Dividing the flow field into piece-wise constant partitions results in a difference between the partitioned flow field and the true flow field. However, if the flow field satisfies certain spatial variation conditions, by designing the path tracking controller of the vehicle to cancel the cross-track flow component of the true flow field, the planned path is still feasible for the vehicle to follow [24], [25]. Therefore, the flow partition error will not impact the feasibility of the planned path.

IV. Partitioning flow field into uniform time-varying clusters

In this section, we consider partitioning the flow field into regions of uniform time-varying flow fields. In the situation where the temporal variation of the flow field consists of non-periodic temporal variation, partitioning the domain into time-invariant partitions may introduce significant difference between the true flow and the partitioned flow field. Therefore, in order to let the partitioned flow field better represents the true flow field, we let the partitioned regions remain unchanged, while the partitioned flow in each region changes with respect to time.

We solve this problem in two steps. The first step implements Algorithm 1 to compute K-means clustering in the time-averaged flow field. Given the partitions $\mathcal{R}$, the second step fits a time series model to flow in each of the partitioned regions. The time series model can be obtained either by physical modeling of the flow field, or by time series modeling techniques. We denote the parametric time series model of the α^{th} region as $\mathbf{f}_\alpha(\Theta_\alpha, t)$, where Θ_α denotes the set of unknown parameters of the flow model in the α^{th} region. Then the partitioned flow can be derived by solving the following regression problem

$$\min_{\Theta_\alpha} J_\alpha = \sum_{\mathbf{x} \in R_\alpha} \sum_{t \in T} ||\mathbf{f}_\alpha(\Theta_\alpha, t) - \mathbf{F}(\mathbf{x}, t)||^2. \tag{5}$$

Claim IV.1. *In any partitioned region R_α, the cost function J_α defined in Eq. (5) is equivalent to the following cost function*

$$J'_\alpha = \sum_{t \in T} \Big[n(R_\alpha) ||\mathbf{f}_\alpha(\Theta_\alpha, t) - \phi_\alpha(t)||^2 + \sum_{\mathbf{x} \in R_\alpha} ||\mathbf{F}(\mathbf{x}, t) - \phi_\alpha(t)||^2 \Big],$$

where $\phi_\alpha(t) = \frac{1}{n(R_\alpha)} \sum_{\mathbf{x} \in R_\alpha} \mathbf{F}(\mathbf{x}, t)$ denotes the spatially averaged flow speed in region R_α at time t. $n(\cdot)$ denotes the cardinality of a set.

The proof of Claim IV.1 is exactly the same as the proof to Claim III.1, and thus the detailed proof of the above Claim is omitted in this paper.

Remark IV.1. *Since the second term in J'_α does not depend on Θ_α, which is the parameters to be optimized, the optimal solution of Eq. (5) equals the optimal solution of*

$$\min_{\Theta_\alpha} \sum_{t \in T} ||\mathbf{f}_\alpha(\Theta_\alpha, t) - \phi_\alpha(t)||^2. \tag{6}$$

Therefore, a time-varying model of the uniform flow in each partitioned region can be derived by fitting $\mathbf{f}_\alpha(\Theta_\alpha, t)$ to the spatially averaged flow in the partitioned region.

V. Simulation Results

In this section, we provide two simulations to validate the performance of the proposed flow partition method. Section V-A presents partitioning result of a simulated jet flow. Particle integral lines in both the true flow and the partitioned flow field are presented to demonstrate the similarity of the two fields. Time-optimal path planning using MEJ in partitioned flow field is also presented to show that flow partition reduces the path planning computation cost. Section V-B describes flow field partition of realistic ocean surface flow field, and shows that the flow partition method represents the flow field with reduced amount of information.

A. Simulated flow field

In this example, we consider partitioning of a field $\mathbf{x} = [x_1, x_2], x_1 \in [0, 2], x_2 \in [0, 1]$. In the domain there are three flow components, a spatially varying component, a temporally varying tidal flow component, and an independent identically distributed Gaussian noise. In the region where $x_2 \in (0.65, 1]$, the simulated flow speed is $\mathbf{F}(\mathbf{x}, t) =$

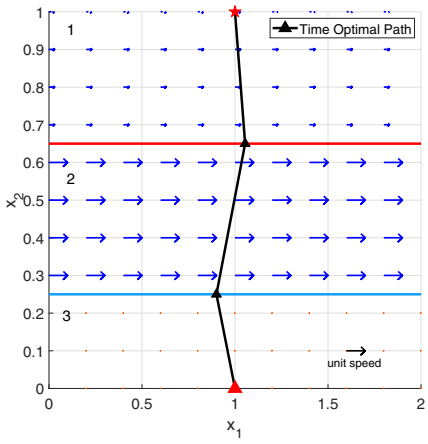

Fig. 1: Partitioned piece-wise constant flow field. Boundaries of partitioned regions are represented by the straight lines. The arrows at each position represent the partitioned flow speed, while the black arrow indicates the arrow length of unit flow speed. The red triangle and star show the vehicle's starting and destination position. The time-optimal path planned by MEJ is represented by the straight line.

$[B_1 + A\sin\omega t + \xi, A\cos\omega t + \xi]^T$. In the region where $x_2 \in [0.25, 0.65], \mathbf{F}(\mathbf{x}, t) = [B_2 + A\sin\omega t + \xi, A\cos\omega t + \xi]^T$, while $\mathbf{F}(\mathbf{x}, t) = [B_3 + A\sin\omega t + \xi, A\cos\omega t + \xi]^T$ in the region where $x_2 \in [0, 0.25)$. We choose the parameters as $B_1 = 0.3, B_2 = 1, B_3 = 0, A = 0.2, \omega = \pi, \xi \sim \mathcal{N}(0, 0.01)$. The flow partition time interval is set to be 10. The vehicle speed is chosen to be 1.5, and the upper bound for maximum flow partition error is set as 0.1.

The flow field partition result is shown in Figure 1. The partitioned flow speed of regions are $\hat{\mathbf{F}}_1 = [0.3022, 0]^T, \hat{\mathbf{F}}_2 = [1.0001, 0]^T, \hat{\mathbf{F}}_3 = [0.0038, 0]^T$. Comparing the partitioned field with the true flow field, the partitioned flow field is a close approximation of the spatially varying component of the original flow field.

We demonstrate path planning using MEJ over the partitioned flow field. The vehicle's starting position is given as $[1, 0]^T$, while the goal position is given as $[1, 1]^T$. The time-optimal path is shown in Figure 1. In order to validate that partitioning the flow field decreases computational cost of path planning, we partition the flow field into increased number of clusters. For each flow partition, we initialize junction positions randomly, and run MEJ five times. Table I shows the averaged MEJ computational time in the flow field partitioned into different number of clusters. As can be seen from the table, the path planning computational cost drops when reducing the number of clusters. Therefore, proper partition of the flow field leads to decreased path planning computation cost.

In order to assess the similarity between the true and the partitioned flow field, we compare the Lagrangian particle trajectories propagated in the true and the partitioned flow fields. 36 particles are initialized uniformly on the domain $x_1 \in [0, 2], x_2 \in [0, 1]$, in both the true flow and partitioned piece-wise constant flow field. The time interval for generating the trajectories is chosen to be 10. Trajectories of the particles are presented in Figure 2. As can be seen from the figure, comparing particle trajectories with the same initial position in the two flow fields, for most of the particles in the domain, the trajectories in the true flow field oscillates near the trajectories in the partitioned flow field. To quantitatively evaluate the difference between particle trajectories, we calculate the RMSE (Root Mean Square Error) between the horizontal and vertical end point difference between the two sets of particle trajectories. The RMSE of end point difference in horizontal direction is 0.5432, approximately 6.8% of the particle displacement in horizontal direction, while the RMSE of end point difference in vertical direction is 0.0754. Therefore, the particle trajectories in the two flow fields are reasonably close. This indicates minor difference between the planned path and the true path.

Number of clusters	Averaged computation time (secs)
3 clusters	0.1110
4 clusters	0.1933
6 clusters	0.2928
8 clusters	0.5284

TABLE I: Computational cost of using MEJ for path planning in the domain of different number of partitions

B. Surface ocean flow field near Gulf Stream

In this example, we present flow field partition result of ocean surface flow field near Cape Hatteras, North Carolina, US. The flow input is given by a 1-km horizontal resolution version of the Navy Coastal Ocean Model (NCOM) [26] made available by J. Book and J. Osborne (Naval Research Laboratory, Stennis Space Center, US). Since the short-term temporal variation of the flow field majorly consists of the tidal flow component, and the lunar semidiurnal M_2 tide is the significant component of tidal flow, the partition time interval is taken over four times the M_2 period (12.42 hours) in order to better decouple the temporal and spatial variation of the field. The 48 hours time-averaged flow field is shown in the left figure in Figure 3.

The Gulf Stream is of relatively high speed, and the flow speed could be comparable to the AUV forward speed. In this case, even small deviation between the true flow and the partitioned flow could potentially cause significant difference between the optimal path in the true flow and the partitioned flow field. Therefore, flow partition should be done in a detailed manner, so that the flow partition error is sufficiently small. We choose the upper bound of flow partition error as 0.3 m/s, approximately 10% of the maximum flow speed in the domain. The middle figure in Figure 3 shows the change of flow partition error with respect to different selection of cluster numbers. Since the flow partition error goes below the upper bound when $k = 11$, the number of clusters is chosen as 11. The partitioned flow field is presented in the right figure in Figure 3.

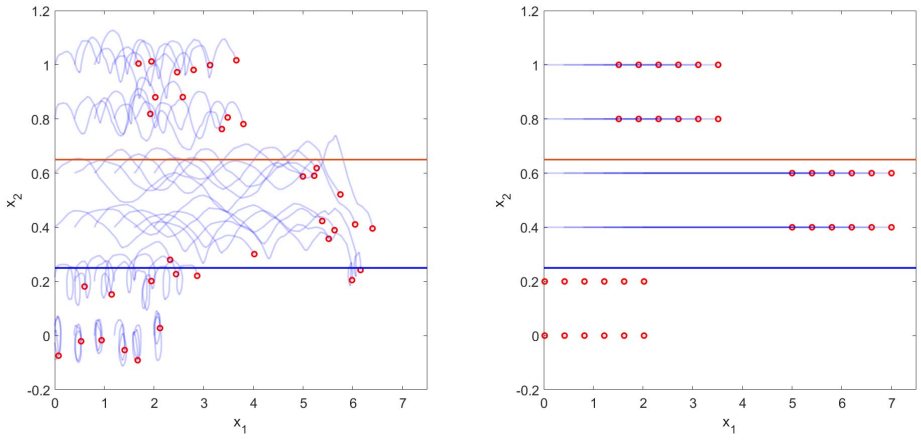

Fig. 2: Left: Trajectories of particles advected by the true flow field. Right: Trajectories of particles advected by the partitioned piece-wise constant flow field. In both figures, the blue curves represent the particle trajectories, while the red circles show particle end positions. The orange and blue lines indicate boundaries of regions.

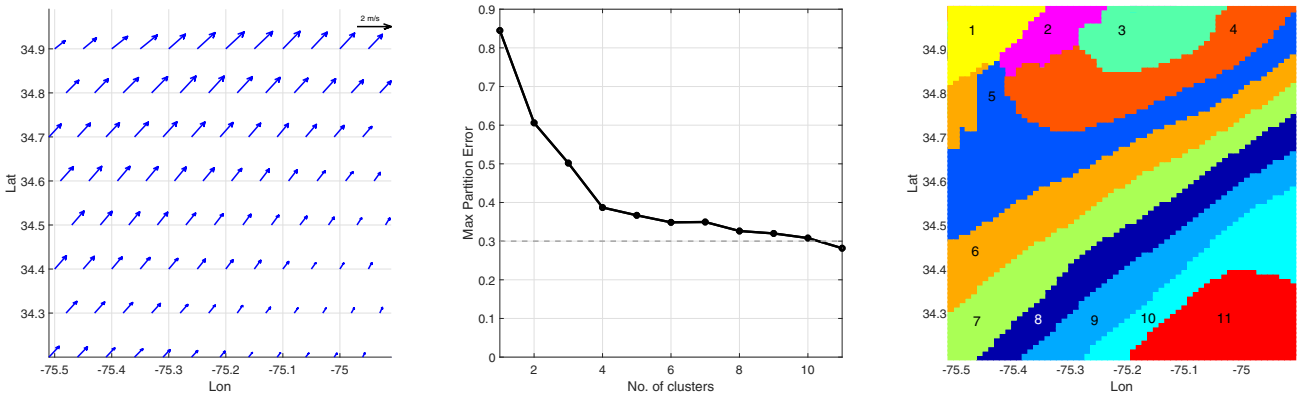

Fig. 3: Left: 48 hours time-averaged flow from May 27, 2017, 00:00 UTC to May 29, 2017, 00:00 UTC at Cape Hatteras, NC. Arrows indicate the flow speed and direction at each grid point in the domain. The black arrow shows reference arrow length of 2 m/s flow. Middle: flow partition error with respect to different number of clusters. Right: Partitioned regions of the flow field.

Due to the complexity of the flow field in the Gulf Stream area, we apply a time series model to better represent the time variation of the flow field. Constrained by the length of the paper, we present the result of extracting temporal variation for only one cluster. We pick region three in the domain, and compute a spatially averaged flow of all flow data in region three during the partition time interval using the ARIMA (Autoregressive Integrated Moving Average) model. The spatially averaged flow is divided into two sets, the first 32 hours data is the training set for identifying the ARIMA model, while the last 16 hours data is the testing set for evaluating performance of the identified model. By applying the Augmented Dickey-Fuller test, and computing the sample autocorrelation and the sample partial autocorrelation of the training data sets, the structure of ARIMA model for both W-E and N-S flow is determined. The parameters of the ARIMA model are then identified using maximum likelihood method. A comparison between the original data points in region three, as well as the spatially averaged data and the predicted spatially averaged data is shown in Figure 4. From the figure, the ARIMA model has better accuracy in predicting the W-E flow component. The maximum W-E flow prediction error is 0.1579 m/s, which is approximately 6.5% of the maximum flow speed in region three. Since the temporal evolution of the training data is more turbulent and shows no clear pattern, the ARIMA model achieves less accuracy in predicting the N-S flow component. The maximum N-S flow prediction error 0.2467 m/s, approximately 10.0% of the maximum flow speed in region three.

The 1-km NCOM surface flow data has approximately 5×10^5 data points in total to represent the dynamic vector field in the domain of interest. With the proposed method, the flow information is reduced to only the flow region boundary data, and the ARIMA model parameters representing both the W-

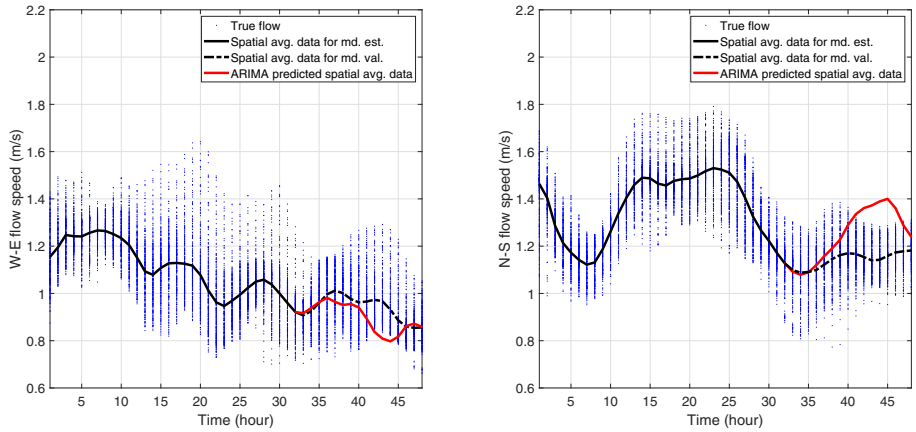

Fig. 4: Comparison between the original flow data, spatially averaged flow data, and the predicted flow data in region 3. The blue dots represent the true flow data at all positions in region 3, binned according to the timestamp of the flow data points. The solid and dashed black curve shows the spatial average of all the data points in region 3 at different time step; the solid curve denotes the data used for estimating the ARIMA model, while the dashed curve is the data used for validating the performance of the identified model. The spatially averaged data predicted by the ARIMA model is represented by the red curve. Left: W-E direction flow comparison. Right: N-S direction flow comparison.

E and N-S flow temporal variation in each region. Therefore, the partitioned flow map significantly reduces the amount of information needed to represent the flow field.

VI. Conclusion

In this paper, we propose a K-means based method to partition a dynamic flow field into static partitions of uniform flow. In the situation where the temporal variation of the flow field only contains a periodic component, we propose to partition the flow field into static clusters of constant flow. In the situation where the flow field exhibits more complex temporal variation, we further extend the proposed method by fitting the spatially averaged flow in each static partition to a parametric time series model. Simulation results demonstrate that the proposed method is a close approximation of the true flow field, while at the same time the path planning computational cost can be reduced by performing path planning over the partitioned flow field. Besides, we show that the proposed method can significantly reduce the amount of data needed to represent the dynamic flow field.

References

[1] N. E. Leonard, D. A. Paley, R. E. Davis, D. M. Fratantoni, F. Lekien, and F. Zhang, "Coordinated control of an underwater glider fleet in an adaptive ocean sampling field experiment in Monterey Bay," *Journal of Field Robotics*, vol. 27, no. 6, pp. 718–740, nov 2010. [Online]. Available: http://doi.wiley.com/10.1002/rob.20366

[2] R. N. Smith, Y. Chao, P. P. Li, D. A. Caron, B. H. Jones, and G. S. Sukhatme, "Planning and implementing trajectories for autonomous underwater vehicles to track evolving ocean processes based on predictions from a Regional Ocean Model," *The International Journal of Robotics Research*, vol. 29, no. 12, pp. 1475–1497, 2010.

[3] P. Ozog, N. Carlevaris-Bianco, A. Y. Kim, and R. M. Eustice, "Long-term mapping techniques for ship hull inspection and surveillance using an autonomous underwater vehicle," *Journal of Field Robotics*, vol. 33, pp. 265–289, 2016.

[4] D. A. Paley, "Cooperative control of collective motion for ocean sampling with autonomous vehicles," Ph. D. thesis, Princeton University, 2007.

[5] F. Zhang, D. M. Fratantoni, D. A. Paley, J. M. Lund, and N. E. Leonard, "Control of coordinated patterns for ocean sampling," *International Journal of Control*, vol. 80, no. 7, pp. 1186–1199, 2007.

[6] F. Zhang, G. Marani, R. N. Smith, and H. T. Choi, "Future trends in marine robotics," *IEEE Robotics and Automation Magazine*, vol. 22, no. 1, pp. 14–21, pp. 122, 2015.

[7] E. Fiorelli, N. E. Leonard, P. Bhatta, D. A. Paley, R. Bachmayer, and D. M. Fratantoni, "Multi-AUV control and adaptive sampling in Monterey Bay," *IEEE Journal of Oceanic Engineering*, vol. 31, no. 4, pp. 935–948, 2006.

[8] S. Liu, J. Sun, J. Yu, A. Zhang, and F. Zhang, "Distributed traversability analysis of flow field under communication constraints," *IEEE Journal of Oceanic Engineering*, 2018.

[9] J. Park, "A reduced element map representation and applications: Map merging, path planning, and target interception," Ph.D. dissertation, Auburn University, 2017.

[10] W. Li, J. Lu, H. Zhou, and S.-N. Chow, "Method of evolving junctions: A new approach to optimal control with constraints," *Automatica*, vol. 78, pp. 72–78, 2017.

[11] H. Zhai, M. Hou, F. Zhang, and H. Zhou, "Method of evolving junction on optimal path planning in flow fields," preprint on webpage at http://arxiv.org/abs/1904.11554.

[12] S.-N. Chow, T.-S. Yang, and H.-M. Zhou, "Global optimizations by intermittent diffusion," in *Chaos, CNN, Memristors and Beyond: A Festschrift for Leon Chua With DVD-ROM, composed by Eleonora Bilotta*. World Scientific, 2013, pp. 466–479.

[13] T. Salzbrunn, H. Leitte, T. Wischgoll, and G. Scheuermann, "The state of the art in flow visualization: Partition-based techniques." vol. 2008, 01 2008, pp. 75–92.

[14] B. Heckel, G. Weber, B. Hamann, and K. I. Joy, "Construction of vector field hierarchies," in *Proceedings of the conference on Visualization'99: celebrating ten years*. IEEE Computer Society Press, 1999, pp. 19–25.

[15] A. Telea and J. J. Van Wijk, "Simplified representation of vector fields," in *Proceedings of the conference on Visualization'99: celebrating ten years*. IEEE Computer Society Press, 1999, pp. 35–42.

[16] Q. Du and X. Wang, "Centroidal voronoi tessellation based algorithms for vector fields visualization and segmentation," in *Proceedings of the conference on Visualization'04*. IEEE Computer Society, 2004, pp. 43–50.

[17] L. Zhang, R. S. Laramee, D. Thompson, A. Sescu, and G. Chen, "An in-

tegral curve attribute based flow segmentation," *Journal of Visualization*, vol. 19, no. 3, pp. 423–436, 2016.
[18] N. Omata and S. Shirayama, "A novel method for unsteady flow field segmentation based on stochastic similarity of direction," *AIP Advances*, vol. 8, no. 4, p. 045020, 2018.
[19] A. Fabregat, I. Mezic, and A. C. Poje, "Finite-time partitions for lagrangian structure identification in Gulf Stream eddy transport," *arXiv preprint arXiv:1606.07382*, 2016.
[20] A. Hadjighasem, D. Karrasch, H. Teramoto, and G. Haller, "Spectral-clustering approach to lagrangian vortex detection," *Physical Review E*, vol. 93, no. 6, p. 063107, 2016.
[21] A. Ramos, V. García-Garrido, A. Mancho, S. Wiggins, J. Coca, S. Glenn, O. Schofield, J. Kohut, D. Aragon, J. Kerfoot *et al.*, "Lagrangian coherent structure assisted path planning for transoceanic autonomous underwater vehicle missions," *Scientific reports*, vol. 8, no. 1, p. 4575, 2018.
[22] G. Gan, C. Ma, and J. Wu, *Data clustering: theory, algorithms, and applications*. Siam, 2007, vol. 20.
[23] G. L. Pickard and W. J. Emery, *Descriptive physical oceanography: an introduction*. Elsevier, 2016.
[24] M. Hou, S. Liu, F. Zhang, and C. R. Edwards, "A combined path planning and path following method for underwater glider navigation in a strong, dynamic flow field," in *OCEANS'18 Kobe, Japan*, 2018.
[25] ——, "Path tracking error analysis for underwater glider navigation in a spatially and temporally varying flow field," in *OCEANS'18 Charleston, SC*, 2018.
[26] P. J. Martin, "Description of the Navy Coastal Ocean Model version 1.0," Naval Research Lab, Tech. Rep. NRL/FR/7322–00-9962, 2000.