Manuscript submitted to
AIMS' Journals
Volume **X**, Number **0X**, XX **200X**

doi:10.3934/xx.xxxxxxx

pp. **X–XX**

MOTION TOMOGRAPHY VIA OCCUPATION KERNELS

Benjamin P. Russo[1*§]

[1] Oak Ridge National Laboratory, Computer Science and Mathematics Division

Rushikesh Kamalapurkar[2,†,‡], Dongsik Chang[3,⋆], and Joel A. Rosenfeld[4,†,‡,**]

[2] Oklahoma State University, School of Mechanical and Aerospace Engineering
[3] Oregon State University, Collaborative Robotics and Intelligent Systems Institute
[4] University of South Florida, Department of Mathematics and Statistics

(Communicated by the associate editor name)

Abstract. The goal of motion tomography is to recover a description of a vector flow field using measurements along the trajectory of a sensing unit. In this paper, we develop a predictor corrector algorithm designed to recover vector flow fields from trajectory data with the use of occupation kernels developed by Rosenfeld et al. [9, 10]. Specifically, we use the occupation kernels as an adaptive basis; that is, the trajectories defining our occupation kernels are iteratively updated to improve the estimation in the next stage. Initial estimates are established, then under mild assumptions, such as relatively straight trajectories, convergence is proven using the Contraction Mapping Theorem. We then compare the developed method with the established method by Chang et al. [5] by defining a set of error metrics. We found that for simulated data, where a ground truth is available, our method offers a marked improvement over [5]. For a real-world example, where ground truth is not available, our results are similar results to the established method.

1. **Introduction.** Over the past decade, unmanned aircraft and underwater systems have evolved significantly and are on the verge of becoming a ubiquitous part of urban and littoral landscape. To compensate for the lack of access to the global positioning system (GPS), unmanned underwater vehicles (UUVs) often rely on the

2020 *Mathematics Subject Classification.* Primary: 93-08; Secondary: 46E22.

Key words and phrases. Motion Tomography, Occupation Kernels, Reproducing Kernel Hilbert Spaces, System Identification.

† Research supported by AFOSR Awards FA9550-20-1-0127 and FA9550-21-1-0134.

‡ Research supported by NSF grants ECCS-2027976, ECCS-2027999 and NRI 2.0 - 1925147.

** The following YouTube playlist discusses the contents of this manuscript: `https://www.youtube.com/playlist?list=PLldiDnQu2phuBWvO4OeiqNve6-_GOdSsz`.

⋆This work was partially completed while the author was at the University of Michigan.

* Corresponding author: russobp@ornl.gov.

§ This work was partially completed while the author was at Farmingdale State College SUNY.

knowledge of the flow field to improve localization [8]. Similarly, accurate estimation of urban wind fields is widely acknowledged to be a significant challenge for implementation of traffic control systems (such as [2]) for unmanned air vehicles (UAVs) [11]. While it is possible to model air flow fields and ocean currents using measurements from on-board sensors, lack of accurate localization (in the case if UUVs) and vechicle-induced noise (in the case of UAVs, especially multi-rotor UAVs), creates significant challenges in acquisition and processing of the data generated by on-board sensors. The aforementioned challenges, along with the payload reduction associated with removing flow sensors, motivates the development of estimation techniques that rely only on the effect of the flow field on the motion of UAVs and UUVs, and not on direct measurements of the flow velocities.

Motion tomography refers to the reconstruction of a vector field using its accumulated effects on mobile sensing units as they travel through the field [14, 5]. Motion tomography allows for the use of low cost mobile underwater/air vehicles as sensors to accumulate sufficient data for estimation of vector fields resulting from wind and ocean currents. As a result, military applications such as ocean current mapping for effective navigation of mine countermeasure UUVs in littoral environments, commercial applications such as wind field mapping for navigation of small package delivery UAVs, and disaster response applications such as wind field mapping for the prediction of flame front propagation and smoke spread, stand to benefit from fast and accurate motion tomography. In this paper we propose an algorithm for motion tomography based on occupation kernels developed in [9, 10]. We provide a proof of convergence via the contraction mapping theorem.

The developed approach to motion tomography has several advantages over existing techniques such as [5]. The flow field is approximated here using the occupation kernels as basis functions for approximation, whereas [5] requires a piecewise constant description of the flow field or a parameterization with respect to Gaussian RBFs. Moreover, [5] employs a renormalization routine which imposes limitations on the motion of the mobile sensors. The proposed occupation kernel method avoids the renormalization and does not add further restrictions on the motion. Finally, the representation of the flow field with respect to the occupation kernel basis allows for the application of the approximation abilities of RKHSs, which are exploited in the convergence analysis.

2. **Tools.**

Definition 2.1. A reproducing kernel Hilbert space (RKHS), H, over a set X is a Hilbert space of real valued functions over the set X such that for all $x \in X$ the evaluation functional $E_x g := g(x)$ is bounded.

Remark 1. Since the evaluation functional is bounded (hence continuous) over a reproducing kernel Hilbert space, the Riesz representation theorem guarantees, for all $x \in X$, the existence of a function $K_x \in H$ such that $\langle g, K_x \rangle_H = g(x)$, where $\langle \cdot, \cdot \rangle_H$ is the inner product for H [7, Chapter 1]. The function K_x is called the reproducing kernel function at x, and $K_y(x) = K(x, y) = \langle K_y, K_x \rangle_H$.

Remark 2. Unless indicated otherwise all inner products are taken to be the Hilbert space inner product, i.e. $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_H$. The same is not true for norms, and differing norms will be indicated by subscripts.

Remark 3. Most of this paper is agnostic to the selection of a kernel. By Moore – Aronszajn Theorem (cf. [1]) there is a one to one correspondence between symmetric

positive semi-definite kernels and reproducing kernel Hilbert spaces. Much of the analysis of the paper and the algorithm will work with any kernel. The two primary examples used in the experiments are the Gaussian Kernels, $K(x,y) = e^{-\frac{1}{\mu}\|x-y\|_2}$, and exponential dot product kernels, $K(x,y) = e^{\frac{x\cdot y}{\mu}}$, in which $\mu \in \mathbb{R}$ is a tunable hyperparameter often called the "kernel width" in the case of Gaussians. Only for estimates involving the Gram matrices do we specify to Gaussian kernels, where many results from scattered data interpolation are leveraged (cf. [13]).

Definition 2.2. Let $X \subset \mathbb{R}^n$ be compact, H be a RKHS of continuous functions over X, and $\gamma : [0,T] \to X$ be a bounded measurable trajectory. The functional $g \mapsto \int_0^T g(\gamma(\tau))d\tau$ is bounded, and may be respresented as $\int_0^T g(\gamma(\tau))d\tau = \langle g, \Gamma_\gamma \rangle_H$, for some $\Gamma_\gamma \in H$ by the Riesz representation theorem. The function Γ_γ is called the occupation kernel corresponding to γ in H.

The value of an inner product against an occupation kernel in a RKHS can be approximated by leveraging quadrature techniques for integration. The numerical experiments described below utilize Simpson's rule when computing inner products involving occupation kernels and their associated Gram matrices, pivotal in the analysis contained in Theorem 4.1. We present Theorem 2.3 which shows convergence with Simpson's Rule. Moreover, the occupation kernels themselves can be expressed as an integral against the kernel function in a RKHS as demonstrated in Proposition 1. Theorem 2.3 and Proposition 1 originally appear in [10], but are reproduced below for completeness.

Proposition 1. *Let H be a RKHS over a compact set X consisting of continuous functions with kernel $K(x,y)$ and let $\gamma : [0,T] \to X$ be a continuous trajectory as in Definition 2.2. The occupation kernel corresponding to γ in H, Γ_γ, may be expressed as*

$$\Gamma_\gamma(x) = \int_0^T K(x, \gamma(t))dt. \tag{1}$$

Proof. Note that $\Gamma_\gamma(x) = \langle \Gamma_\gamma, K(\cdot, x)\rangle_H$, by the reproducing property of K. Consequently,

$$\Gamma_\gamma(x) = \langle \Gamma_\gamma, K(\cdot,x)\rangle_H = \langle K(\cdot,x), \Gamma_\gamma\rangle_H = \int_0^T K(\gamma(t),x)dt = \int_0^T K(x,\gamma(t))dt,$$

which establishes the result. □

Leveraging Proposition 1, quadrature techniques can be demonstrated to give not only pointwise convergence but also norm convergence in the RKHS, which is a strictly stronger result.

Theorem 2.3. *Under the hypothesis of Proposition 1, let $t_0 = 0 < t_1 < t_2 < \ldots < t_F = T$ (with F even and t_i evenly spaced), suppose that γ is a fourth order continuously differentiable trajectory and H is composed of fourth order continuously differentiable functions. Set h to satisfy $t_i = t_0 + ih$, and consider*

$$\hat{\Gamma}_\gamma(x) := \frac{h}{3}\left(K(x,\gamma(t_0)) + 4\sum_{i=1}^{\frac{F}{2}} K(x,\gamma(t_{2\cdot i-1})) + 2\sum_{i=1}^{\frac{F}{2}-1} K(x,\gamma(t_{2\cdot i})) + K(x,\gamma(t_F))\right). \tag{2}$$

The norm distance is bounded as $\|\Gamma_\gamma - \hat{\Gamma}_\gamma\|_H^2 = O(h^4)$.

Proof. Consider $\|\Gamma_\gamma - \hat{\Gamma}_\gamma\|_H^2 = \langle \Gamma_\gamma, \Gamma_\gamma \rangle_H + \langle \hat{\Gamma}_\gamma, \hat{\Gamma}_\gamma \rangle_H - 2\langle \Gamma_\gamma, \hat{\Gamma}_\gamma \rangle_H$. The term $\langle \hat{\Gamma}_\gamma, \hat{\Gamma}_\gamma \rangle_H$ is an implementation of the two-dimensional Simpson's rule (cf. [4]) while $\langle \Gamma_\gamma, \Gamma_\gamma \rangle_H$ is the double integral $\int_0^T \int_0^T K(\gamma(t), \gamma(\tau)) dt d\tau$. Thus,

$$\langle \hat{\Gamma}_\gamma, \hat{\Gamma}_\gamma \rangle_H = \langle \Gamma_\gamma, \Gamma_\gamma \rangle_H + O(h^4).$$

Similarly, $\langle \Gamma_\gamma, \hat{\Gamma}_\gamma \rangle_H$ integrates in one variable while implementing Simpson's rule in the other. Consequently,

$$\langle \Gamma_\gamma, \hat{\Gamma}_\gamma \rangle_H = \langle \Gamma_\gamma, \Gamma_\gamma \rangle_H + O(h^4).$$

The conclusion of the theorem follows. □

It should be noted that the above convergence rate is for norm convergence and point-wise convergence can be demonstrated to be faster [10].

3. **Problem Setup and Developed Algorithm.** Let $r : [0, T] \to \mathbb{R}^2$ represent a continuous trajectory for a mobile sensor attempting to travel in a straight line, but subject to an unknown flow field, $F : X \subset \mathbb{R}^2 \to \mathbb{R}^2$, where X is a compact subset of the plane. Let $\dot{r} = s \begin{pmatrix} \cos(\theta) & \sin(\theta) \end{pmatrix}^T + F(r)$, for a positive constant s, represent the true dynamics induced by the flow field. We will assume $F : X \to \mathbb{R}^2$ is locally Lipschitz in order to assure uniqueness of the solutions [12]. As the flow field is unknown, the anticipated dynamics are given as $\dot{\tilde{r}} = s \begin{pmatrix} \cos(\theta) & \sin(\theta) \end{pmatrix}^T$. After a mobile sensor has traveled through the flow field over a time period $[0, T]$, during which there is limited to no knowledge of the mobile sensor's position, the difference between the actual location of the mobile sensor, $r(T)$, and the anticipated location, $\tilde{r}(T)$, is given as

$$D = r(T) - \tilde{r}(T) = \int_0^T \left(\dot{r}(t) - \dot{\tilde{r}}(t) \right) dt = \int_0^T F(r(t)) dt$$

Hence, $D = \langle F, \Gamma_r \rangle_H$ and the difference between the final locations of the mobile sensor describes the integral of the flow field along the trajectory r. This integral provides a type of tomographic sample of F. As the trajectory r is treated as unknown, an approximation of F using the sample generated by Γ_r is difficult to assess directly. This motivates the developed iterative algorithm to determine the flow field, F, as well as the true trajectories, r. On each iteration, we use the current approximation to calculate a simulated trajectory. The difference between the end points of the simulated and the real trajectory is then used to update our vector-field through solving a matrix equation. Using notation to be established below: For each $i = 1, \ldots, s$,

$$w_{1,n} \int_0^T \Gamma_{\tilde{r}_{1,n}}(\tilde{r}_{i,n}(t)) dt + \cdots + w_{s,n} \int_0^T \Gamma_{\tilde{r}_{s,n}}(\tilde{r}_{i,n}(t)) dt = D_{i,n} + \int_0^T \hat{F}_n(\tilde{r}_{i,n}(t)) dt.$$

In this form, we see that the $D_{i,n}$ acts as a correction term.

Let $\{s_i\}_{i=1}^M$ and $\{\theta_i\}_{i=1}^M$ be a collection of speeds and angles used to generate a collection of anticipated trajectories, i.e. trajectories governed by the dynamics $\dot{\tilde{r}} = s_i \begin{pmatrix} \cos(\theta_i) & \sin(\theta_i) \end{pmatrix}^\top$. Moreover, let $\{p_i\}_{i=1}^M \subset \mathbb{R}^2$ represent the starting point of the trajectories.

Significantly, after the initial data collection period, no further experiments are necessary to approximate the flow field. Specifically, Algorithm 1 only needs to produce new simulations of the approximate trajectories which ultimately converge to the true trajectories from the initial experiment.

Algorithm 1 Iterative Motion Tomography Algorithm

Define N as the number of iterates
Input: Samples $r_i(T) \quad i \in \{1, \dots, s\}$
Set $D_{i,0} = r_i(T) - \tilde{r}_{i,0}(T) \quad i \in \{1, \dots, s\}$
Set $\hat{F}_0 = 0$
for $n \in \{0, \dots, N\}$ **do**
  for $i \in \{1, \dots, s\}$ **do**
    Generate, via a numerical method, $\tilde{r}_{i,n} : [0, 1] \to X$ as the unique solution to

$$\dot{p} = s_i \begin{pmatrix} \cos(\theta_i) & \sin(\theta_i) \end{pmatrix}^\top + \hat{F}_n(p), \quad p(0) = p_i.$$

  end for
  Set $D_{i,n} = r_i(T) - \tilde{r}_{i,n}(T) \quad i \in \{1, \dots, s\}$
  for $i \in \{1, \dots, s\}$ **do**
    Compute $\{w_{i,n+1}\}_{i=1}^s$ by solving

$$\left(\langle \Gamma_{\tilde{r}_{i,n}}, \Gamma_{\tilde{r}_{j,n}} \rangle\right)_{i,j=1}^{s,s} \begin{pmatrix} w_{1,n+1} \\ \vdots \\ w_{s,n+1} \end{pmatrix} = \begin{pmatrix} D_{1,n} + \langle \hat{F}_n, \Gamma_{\tilde{r}_{1,n}} \rangle \\ \vdots \\ D_{M,n} + \langle \hat{F}_n, \Gamma_{\tilde{r}_{s,n}} \rangle \end{pmatrix}.$$

  end for
  Output $\hat{F}_{n+1} = \sum_{i=1}^s w_{i,n+1} \Gamma_{\tilde{r}_{i,n}}$
end for

4. **Convergence of Algorithm.** This section is devoted to establishing sufficient conditions for the above algorithm to converge. The main theoretical crux will be the contraction mapping theorem. Let $X \subset \mathbb{R}^2$ be compact, $F : X \to \mathbb{R}^2$ a Lipschitz continuous vector field, let $\{r_i\}_{i=1}^\infty$ be a countable set of trajectories, $r_i : [0, T] \to X$, satisfying

$$\dot{p} = s_i \begin{pmatrix} \cos(\theta_i) & \sin(\theta_i) \end{pmatrix}^\top + F(p), \quad p(0) = p_i$$

for some countable collection of $(s_i, \theta_i) \subset \mathbb{R} \times [0, 2\pi)$, and countable set $\{p_i\}_{i=1}^\infty \subset X$. Let $H(X)$ be a reproducing kernel Hilbert space of $\mathbb{R}^2$ valued functions.

Theorem 4.1. *Let $\Omega \subset H(X)$ be the subset of $\phi(x) = \begin{pmatrix} \phi_1(x) & \phi_2(x) \end{pmatrix}^\top$ with $\phi_i(x) = \sum_{j=1}^s w_{j,\phi}^i \Gamma_{r_{j,\phi}}(x)$ for $i = 1, 2$. We define $A : \Omega \subset H(X) \to H(X)$ as follows: Given, $\phi \in \Omega$ define $\hat{r}_{j,\phi}$ to be the solution to*

$$\dot{p} = s_j \begin{pmatrix} \cos(\theta_j) & \sin(\theta_j) \end{pmatrix}^\top + \phi(p) \quad p(0) = p_j.$$

Moreover, define $D_{j,\phi}^i := r_j^i(T) - \hat{r}_{j,\phi}^i(T)$ for $i = 1, 2$. Then

$$A(\phi)_i := \sum_{j=1}^s \hat{w}_{j,\phi}^i \Gamma_{\hat{r}_{j,\phi}}(x)$$

where $\hat{w}_{j,\phi}^i$ are defined so that $A(\phi)_i$ satisfies

$$\langle A(\phi)_i, \Gamma_{\hat{r}_{j,\phi}} \rangle_H = D_{j,\phi}^i + \langle \phi, \Gamma_{\hat{r}_{j,\phi}} \rangle_H$$

for $i = 1, 2$. Under the set of assumptions, (1) – (2), listed in following section there exists a closed finite diameter subspace $E \subset \Omega$ such that $A|_E$ is a contraction mapping and thus extends to a contraction mapping on $H(X)$.

We note that Hilbert spaces have the Lipschitz extension property [3], i.e. a Banach space E has the Lipschitz (Contraction) extension property if every Lipschitz map on a subset of E extended to all of E with the same Lipschitz constant. We will be proving that the mapping described above is a contraction in some ball of diameter less than one containing the solution and applying a version of Contraction Mapping Theorem (cf. [6]) to prove convergence. Furthermore, we will also be suppressing the index $i = 1, 2$ and the use of operator valued kernels to facilitate the simplicity of presentation, and it should be noted that the proof extends to any finite dimension. The operator valued kernels being used are of the form $\mathbb{K}(x, y) = \text{diag}(K(x, y))$.

Theorem 4.2 (Contraction Mapping Theorem)**.** *Let X be a complete metric space and let S_0 be a closed subset of X of finite diameter. Let $P : S_0 \to S_0$ be a contraction mapping. Then the sequence of iterates $\{x_k\}$ produced by successive iterations $x_{k+1} = P(x_k)$ converges to $x = P(x)$, the unique fixed point of P in S_0 for any x_0 in S_0.*

Remark 4. The contraction mapping theorem guarantees convergence in the metric. As the metric space is a reproducing kernel Hilbert space the metric is given as a norm difference. Hence, the convergence is in terms of the norm, which for reproducing kernel Hilbert spaces also gives point-wise convergence.

To prove Theorem 4.1 we establish several propositions that give inequalities relating norms to the norms of the flow fields.

Proposition 2. *Let $H(X)$ be the reproducing kernel Hilbert space described above. For $\phi, \psi \in H(X)$, if r_ϕ and r_ψ are the unique solutions to the initial value problems*

$$\dot{p} = s\begin{pmatrix}\cos(\theta) & \sin(\theta)\end{pmatrix}^\top + \phi(p), \quad p(0) = p$$

$$\dot{p} = s\begin{pmatrix}\cos(\theta) & \sin(\theta)\end{pmatrix}^\top + \psi(p), \quad p(0) = p \in X$$

respectively, then

$$|r_\phi(t) - r_\psi(t)| \leq M\|\phi - \psi\|_H$$

for some constant M.

Proof. We have that

$$\begin{aligned}|\dot{r}_\phi - \dot{r}_\psi| = |\phi(r_\phi(t)) - \psi(r_\psi(t))| &\leq |\phi(r_\phi(t)) - \phi(r_\psi(t))| + |\phi(r_\psi(t)) - \psi(r_\psi(t))| \\ &\leq \|\nabla\phi\|_\infty \cdot |r_\phi(t) - r_\psi(t)| \\ &\quad + \|\phi - \psi\|_H\sqrt{K(r_\psi(t), r_\psi(t))}\end{aligned}$$

by an application of mean value theorem. Note that $\|\cdot\|_\infty$ denotes the sup-norm. It follows that,

$$|r_\phi(t) - r_\psi(t)| \leq \exp(\|\nabla\phi\|_\infty)|r_\phi(0) - r_\psi(0)| + \|\phi - \psi\|_H \cdot M = \|\phi - \psi\|_H \cdot M$$

for $M = \max_x K(x, x)$ which exists by compactness of the region. □

The proof of the following proposition can be found in [10], but is presented here for clarity.

Proposition 3. *Suppose H is a RKHS over a set X consisting of continuous functions and let $\gamma_1(t)$ and $\gamma_2(t)$ be two trajectories with homotopy $\{\gamma_w(t)\}$. The map $w \mapsto \Gamma_{\gamma_w}$ is continuous.*

Proof. As $[0,T]\times[0,1]$ is compact, the map $(t,w)\mapsto\gamma_w(t)$ is uniformly continuous. That is for every $\varepsilon>0$ there exists a $\delta>0$ such that whenever $\|(t_1,w_1)-(t_2,w_2)\|_2<\delta$, $\|\gamma_{w_1}(t_1)-\gamma_{w_2}(t_2)\|_2<\epsilon$. Moreover, as $K(\cdot,\cdot)$ is continuous, and the image of $\gamma_w(t)$ is compact, the map $(w_1,t,w_2,\tau)\mapsto K(\gamma_{w_1}(t),\gamma_{w_2}(\tau))$ is uniformly continuous.

Fix $\varepsilon>0$ and select $\delta>0$ such that

$$|K(\gamma_{w_1}(t),\gamma_{w_1}(\tau))-K(\gamma_{w_2}(t),\gamma_{w_1}(\tau))|<\frac{\varepsilon}{2T^2}\quad\text{and}$$
$$|K(\gamma_{w_2}(t),\gamma_{w_2}(\tau))-K(\gamma_{w_2}(t),\gamma_{w_1}(\tau))|<\frac{\varepsilon}{2T^2}$$

whenever $|w_1-w_2|<\delta$. Select w_1,w_2 such that $|w_1-w_2|<\delta$, then

$$\begin{aligned}\|\Gamma_{w_1}-\Gamma_{w_2}\|_H^2&=\langle\Gamma_{w_1},\Gamma_{w_1}\rangle+\langle\Gamma_{w_2},\Gamma_{w_2}\rangle-2\langle\Gamma_{w_1},\Gamma_{w_2}\rangle\\&=\int_0^T\int_0^T(K(\gamma_{w_1}(t),\gamma_{w_1}(\tau))-K(\gamma_{w_2}(t),\gamma_{w_1}(\tau))dtd\tau\\&+\int_0^T\int_0^T K(\gamma_{w_2}(t),\gamma_{w_2}(\tau))-K(\gamma_{w_2}(t),\gamma_{w_1}(\tau)))dtd\tau,\end{aligned}\tag{3}$$

Note that Equation (3) is positive and bounded by ε by construction. Hence, the map $w\mapsto\Gamma_w$ is continuous. □

Corollary 1. *Following the notational convention defined in Proposition 3, if $|r_\phi(t)-r_\psi(t)|\le M\|\phi-\psi\|_H$ for some M then $\|\Gamma_{r_\phi}-\Gamma_{r_\psi}\|_H\le\sqrt{C}T\|\phi-\psi\|_H^{1/2}$ for some constant C independent of T.*

Proof.

$$\begin{aligned}\|\Gamma_{r_\phi}-\Gamma_{r_\psi}\|_H^2&=\langle\Gamma_{r_\phi},\Gamma_{r_\phi}\rangle+\langle\Gamma_{r_\psi},\Gamma_{r_\psi}\rangle-2\langle\Gamma_{r_\phi},\Gamma_{r_\psi}\rangle\\&=\int_0^T\int_0^T K(r_\phi(t),r_\phi(\tau))-K(r_\psi(t),r_\phi(\tau))\,dt\,d\tau\\&+\int_0^T\int_0^T K(r_\psi(t),r_\psi(\tau))-K(r_\psi(t),r_\phi(\tau))\,dt\,d\tau\\&\le\int_0^T\int_0^T(|\nabla_y K(r_\phi(t),\xi_\psi)|)\cdot|r_\phi(\tau)-r_\psi(\tau)|\,dt\,d\tau\\&+\int_0^T\int_0^T(|\nabla_y K(r_\psi(t),\xi_\phi)|)\cdot|r_\phi(\tau)-r_\psi(\tau)|\,dt\,d\tau\\&\le CT^2\|\phi-\psi\|_H.\end{aligned}\tag{4}$$

The last two inequalities are due to an application of mean value theorem and proposition 2. □

In order to talk about the stability of the interpolation process we define the following:

Definition 4.3. The trajectory separation distance is given by

$$q_{X,T}:=\frac{1}{2}\min_{\substack{j\neq k\\ t,\tau\in[0,T]}}\|\gamma_j(t)-\gamma_k(\tau)\|_2$$

This is the maximum radius such that all "tubes" centered on the trajectories are disjoint.

The next theorem will be a trajectory variant of a theorem found in Wendland [13]. In Wendland, a bound for the minimum eigenvalue is obtained by expressing the kernel function $\Phi(x-y)$ in terms of its Fourier transform and comparing with an intermediate radial functions Ψ_M based on characteristic functions for balls of radius M. This comparison is given in terms of the separation distance between centers. By defining the trajectory separation distance as above the inequalities found in Wendland remain valid since the points in the image of each trajectory are separated by at least $q_{X,T}$.

In the following theorem let G_Γ be the Occupation kernel Grammian, i.e. for a set of trajectories, $r_i : [0,T] \to \mathbb{R}^2$, with $1 \leq i \leq N$

$$G_\Gamma := \left(\langle \Gamma_{r_j}, \Gamma_{r_k} \rangle\right)_{j,k=1}^N = \int_0^T \int_0^T \left(K(r_j(t), r_k(\tau))\right)_{j,k=1}^N \, dt \, d\tau.$$

We also need the following definition found in [13]:

Definition 4.4. We say a function f is slowly increasing if there exists a constant $m \in \mathbb{N}_0$ such that $f(x) = \mathcal{O}(\|x\|_2^m)$ for $\|x\|_2 \to \infty$.

Remark 5. Essentially a function is "slowly increasing" if it is polynomially bounded for large x. As Gaussians decay to zero as $\|x\|_2 \to \infty$, Gaussians are bounded by a constant and fall under this category.

Theorem 4.5. *With $\Phi(\|x-y\|_2) = K(x,y) = \exp(-\mu\|x-y\|_2^2)$ the minimal eigenvalue of the Occupation kernel Grammian is bounded by*

$$\lambda_{\min(G_\Gamma)} \geq \frac{C_2}{2\mu} \frac{\exp(-M_2^2/(q_{X,T}^2\mu))T^2}{q_{X,T}^2}$$

with $M_2 = 12\sqrt[3]{\frac{\pi}{9}}$ and $C_2 = \frac{M^2}{16}$.

Proof. Let $\Phi : \mathbb{R}^d \to \mathbb{C}$ be any radial, continuous, and slowly increasing function possessing a non-negative generalized Fourier transform linearity, for $\alpha = (\alpha_1, \ldots, \alpha_s)^\top$, given,

$$\alpha^\top G_{\Phi(r)} \alpha = \sum_{j,k=1}^s \alpha_j \alpha_k \Phi(r_j(t) - r_k(\tau)) = \int_{\mathbb{R}^d} \sum_{j,k=1}^s \alpha_j \alpha_k e^{i\omega^\top (r_j(t) - r_k(t))} \hat{\Phi}(\omega) \, d\omega$$

by linearity an application of Fubini we have,

$$\begin{aligned} \alpha^\top G_\Gamma \alpha &= \int_0^T \int_0^T \alpha^\top G_{\Phi(r)} \alpha \, dt \, d\tau \qquad (5) \\ &= \int_0^T \int_0^T \int_{\mathbb{R}^d} \sum_{j,k=1}^s \alpha_j \alpha_k e^{i\omega^\top (r_j(t) - r_k(\tau))} \hat{\Phi}(\omega) \, d\omega \\ &= \int_{\mathbb{R}^d} \left| \sum_{j=1}^s \alpha_j \int_0^T e^{i\omega^\top r_j(t)} \, dt \right|^2 \hat{\Phi}(\omega) \, d\omega \end{aligned}$$

where (5) is valid whenever $\Phi : \mathbb{R}^d \to \mathbb{C}$ is a continuous and slowly increasing function possessing a non-negative generalized Fourier transform. Define,

$$\psi_{M_2}(x) = \frac{\varphi_0(M_2)\Gamma\left(\frac{d}{2}+1\right)}{2^{d/2}} \|x\|_2^{-d} J_{d/2}^2(M_2\|x\|_2)$$

where J_ν is the Bessel function of the first kind and

$$\varphi_0(M_2) = \inf_{\|\omega\|_2 \leq 2M_2} \hat{\Phi}(\omega).$$

In Wendland [13] it is established that for a collection of points x_i, $i = 1, \dots s$,

$$\begin{aligned}\int_{\mathbb{R}^d} \sum_{j,k=1}^{s} \alpha_j \alpha_k e^{i\omega^\top (x_j - x_k)} \hat{\psi}_{M_2}(\omega)\, d\omega &= \sum_{j,k=1}^{s} \alpha_j \alpha_k \psi_{M_2}(x_j - x_k) \\ &\geq \|\alpha\|_2^2 \left(\psi_{M_2}(0) - \max_{1 \leq j \leq s} \sum_{k=1, k\neq j}^{s} |\psi_{M_2}(x_j - x_k)| \right)\end{aligned}$$

and that for the given value of M_2,

$$\max_{1 \leq j \leq s} \sum_{k=1, k\neq j}^{s} |\psi_{M_2}(x_j - x_k)| \leq \frac{1}{2}\psi_{M_2}(0).$$

Thus, for the stated value of M_2 we have that,

$$\sum_{j,k=1}^{s} \alpha_j \alpha_k \psi_{M_2}(x_j - x_k) \geq \|\alpha\|_2^2 \frac{\psi_{M_2}(0)}{2}. \tag{6}$$

Now, the key insight is that the establishment of (6) is unaffected by letting $x_j = r_j(t), x_k = r_k(\tau)$ for $t, \tau \in [0, T]$ by replacing the standard separation distance q_X with $q_{X,T}$ and that this holds over $[0, T]$.

Our auxiliary function ψ_{M_2} is designed so that $\hat{\psi}_{M_2}(\omega) \leq \hat{\Phi}(\omega)$, giving us

$$\int_{\mathbb{R}^d} \left| \sum_{j=1}^{s} \alpha_j \int_0^T e^{i\omega^\top r_j(t)}\, dt \right|^2 \hat{\Phi}(\omega)\, d\omega \geq \int_{\mathbb{R}^d} \left| \sum_{j=1}^{s} \alpha_j \int_0^T e^{i\omega^\top r_j(t)}\, dt \right|^2 \hat{\psi}_{M_2}(\omega)\, d\omega. \tag{7}$$

By combining equations, (5), (5), (6), and (7) we get

$$\begin{aligned}\alpha^\top G_\Gamma \alpha = \int_0^T \int_0^T \alpha^\top G_{\Phi(r)} \alpha\, dt\, d\tau &= \int_0^T \int_0^T \int_{\mathbb{R}^d} \sum_{j,k=1}^{s} \alpha_j \alpha_k e^{i\omega^\top (r_j(t) - r_k(\tau))} \hat{\Phi}(\omega)\, d\omega \\ &\geq \int_0^T \int_0^T \|\alpha\|_2^2 \frac{\psi_{M_2}(0)}{2}\, dt\, d\tau.\end{aligned}$$

Finally,

$$\lambda_{\min(G_\Gamma)} = \inf \frac{\alpha^\top G_\Gamma \alpha}{\|\alpha\|_2^2} \geq \frac{C_2}{2\mu} \frac{\exp(-M_2^2/(q_{X,T}^2 \mu)) T^2}{q_{X,T}^2}$$

with $M_2 = 12\sqrt[3]{\frac{\pi}{9}}$ and $C_2 = \frac{M_2^2}{16}$ and using the bounds for $\psi_M(0)$ established in [13]. □

As a corollary to the above analysis we show that the condition number of the interpolation matrices is at least as good as the standard Gaussian kernels.

Corollary 2. *If*

$$G_\Gamma = \left(\langle \Gamma_{r_j}, \Gamma_{r_k} \rangle \right)_{j,k=1}^{s,s} = \int_0^T \int_0^T \left(K(r_j(t), r_k(\tau)) \right)_{j,k=1}^{s,s} dt\, d\tau,$$

then

$$\lambda_{\max}(G_\Gamma) \leq s T^2 \Phi(0)$$

where $\Phi(\|x_i - x_j\|_2) = K(x_j, x_i)$.

Proof. By a standard application of Gershgorin's circle theorem on eigenvalues,

$$\lambda_{\max}(G_\Gamma) \leq s \max_{i,j=1,\dots s} |\Phi(r_j(t) - r_k(\tau))| \leq s \int_0^T \int_0^T |\Phi(0)|\, dt\, d\tau = sT^2|\Phi(0)|$$

where the above is due to the properties of positive semi-definite functions. □

Since the condition number of an interpolation matrix A is given by the ratio of the maximum and minimum eigenvalues, the presence of the T^2 in the above inequality nullifies the $1/T^2$ appearing in estimate for the minimal eigenvalue. In essence, this shows that we can take the minimal distance between trajectories as a measurement of how good our approximations will be.

4.1. **Main Inequalities.** Following the notion established in statement of Theorem 4.1 consider the following,

$$\begin{aligned}\|A(\phi) - A(\psi)\|_H &= \left\| \sum_{j=1}^s \hat{w}_{j,\phi} \Gamma_{\hat{r}_{j,\phi}}(x) - \sum_{j=1}^s \hat{w}_{j,\psi} \Gamma_{\hat{r}_{j,\psi}}(x) \right\|_H \\ &\leq \sum_{j=1}^s |\hat{w}_{j,\phi}| \|\Gamma_{\hat{r}_{j,\phi}} - \Gamma_{\hat{r}_{j,\psi}}\|_H + \sum_{j=1}^s |\hat{w}_{j,\phi} - \hat{w}_{j,\psi}| \|\Gamma_{\hat{r}_{j,\psi}}\|_H \end{aligned} \tag{8}$$

To proceed forward, we will need to establish control on terms above in as a function of $\|\phi - \psi\|_H$. Note by definition $\begin{pmatrix} \hat{w}_{1,\phi} \cdots \hat{w}_{s,\phi} \end{pmatrix}^\top$ satisfies

$$\begin{pmatrix} \langle \Gamma_{\hat{r}_{1,\phi}}, \Gamma_{\hat{r}_{1,\phi}} \rangle & \cdots & \langle \Gamma_{\hat{r}_{s,\phi}}, \Gamma_{\hat{r}_{1,\phi}} \rangle \\ \vdots & \ddots & \vdots \\ \langle \Gamma_{\hat{r}_{1,\phi}}, \Gamma_{\hat{r}_{M,\phi}} \rangle & \cdots & \langle \Gamma_{\hat{r}_{M,\phi}}, \Gamma_{\hat{r}_{M,\phi}} \rangle \end{pmatrix} \begin{pmatrix} \hat{w}_{1,\phi} \\ \vdots \\ \hat{w}_{s,\phi} \end{pmatrix} = \begin{pmatrix} D_{1,\phi} + \langle \phi, \Gamma_{\hat{r}_{1,\phi}} \rangle \\ \vdots \\ D_{s,\phi} + \langle \phi, \Gamma_{\hat{r}_{s,\phi}} \rangle \end{pmatrix}. \tag{9}$$

For notational convenience, given a $\phi \in H(X)$ let

$$G_\phi = \begin{pmatrix} \langle \Gamma_{\hat{r}_{1,\phi}}, \Gamma_{\hat{r}_{1,\phi}} \rangle & \cdots & \langle \Gamma_{\hat{r}_{s,\phi}}, \Gamma_{\hat{r}_{1,\phi}} \rangle \\ \vdots & \ddots & \vdots \\ \langle \Gamma_{\hat{r}_{1,\phi}}, \Gamma_{\hat{r}_{s,\phi}} \rangle & \cdots & \langle \Gamma_{\hat{r}_{s,\phi}}, \Gamma_{\hat{r}_{s,\phi}} \rangle \end{pmatrix}$$

Now,

$$\left\| \begin{pmatrix} \hat{w}_{1,\phi} \\ \vdots \\ \hat{w}_{s,\phi} \end{pmatrix} - \begin{pmatrix} \hat{w}_{1,\psi} \\ \vdots \\ \hat{w}_{s,\psi} \end{pmatrix} \right\|_2 \leq \|G_\phi^{-1}\|_{\text{op}} \left\| \begin{pmatrix} D_{1,\phi} - D_{1,\psi} + \langle \phi, \Gamma_{\hat{r}_{1,\phi}} \rangle - \langle \psi, \Gamma_{\hat{r}_{1,\psi}} \rangle \\ \vdots \\ D_{s,\phi} - D_{s,\psi} + \langle \phi, \Gamma_{\hat{r}_{s,\phi}} \rangle - \langle \psi, \Gamma_{\hat{r}_{s,\psi}} \rangle \end{pmatrix} \right\|_2 \tag{10}$$

$$+ \|G_\phi^{-1} - G_\psi^{-1}\|_{\text{op}} \left\| \begin{pmatrix} D_{1,\psi} + \langle \psi, \Gamma_{\hat{r}_{1,\psi}} \rangle \\ \vdots \\ D_{s,\psi} + \langle \psi, \Gamma_{\hat{r}_{s,\psi}} \rangle \end{pmatrix} \right\|_2, \tag{11}$$

where $\|\cdot\|_{\text{op}}$ denotes the operator norm induced by the 2-norm. A portion of the proof for Theorem 4.1 will be on establishing suitable control over of the norms in (10), and (11). Before proceeding, we will need the following lemma.

Lemma 4.6. *Given a $\phi, \psi \in H(X)$, then*

$$\|G_\phi - G_\psi\|_{op} \leq ST^2\|\phi - \psi\|_H^{1/2}$$

for a constant $S \propto s$ and independent of T

Proof. Similar to the analysis in corollary 2

$$\begin{aligned}\lambda_{\max}(G_\phi - G_\psi) &\leq s \max_{i,j=1,\dots s} |\langle \Gamma_{\hat{r}_{i,\phi}} - \Gamma_{\hat{r}_{i,\psi}}, \Gamma_{\hat{r}_{j,\phi}} - \Gamma_{\hat{r}_{j,\psi}}\rangle| \\ &\leq s \max_{i,j=1,\dots,s} \|\Gamma_{\hat{r}_{i,\phi}} - \Gamma_{\hat{r}_{i,\psi}}\|_H \cdot \|\Gamma_{\hat{r}_{j,\phi}} - \Gamma_{\hat{r}_{j,\psi}}\|_H \\ &\leq sCT^2\|\phi - \psi\|_H.\end{aligned}$$

Hence, by an application of corollary 1 we have

$$\lambda_{\max}(G_\phi - G_\psi) \leq sCT^2\|\phi - \psi\|_H^{1/2}.$$

Letting $S = sC$ completes our proof. □

Analysis of the upper bounds: Given that by definition

$$|D_{j,\phi} - D_{j,\psi}| = |\hat{r}_{j,\phi}(T) - \hat{r}_{j,\psi}(T)| \leq M \cdot \|\phi - \psi\|_H$$

by Proposition 2 and

$$|\langle \phi, \Gamma_{\hat{r}_{j,\phi}}\rangle - \langle \psi, \Gamma_{\hat{r}_{j,\psi}}\rangle| \leq \|\phi\|_H\|\Gamma_{\hat{r}_{j,\phi}} - \Gamma_{\hat{r}_{j,\psi}}\|_H + \|\phi - \psi\|_H \cdot \|\Gamma_{\hat{r}_{j,\psi}}\|_H.$$

Note that,

$$\left\| \begin{pmatrix} D_{1,\phi} - D_{1,\psi} + \langle \phi, \Gamma_{\hat{r}_{1,\phi}}\rangle - \langle \psi, \Gamma_{\hat{r}_{1,\psi}}\rangle \\ \vdots \\ D_{s,\phi} - D_{s,\psi} + \langle \phi, \Gamma_{\hat{r}_{s,\phi}}\rangle - \langle \psi, \Gamma_{\hat{r}_{s,\psi}}\rangle \end{pmatrix} \right\|_2^2 \leq \sum_{j=1}^{s} |D_{j,\phi} - D_{j,\psi}|^2 + |\langle \phi, \Gamma_{\hat{r}_{j,\phi}}\rangle - \langle \psi, \Gamma_{\hat{r}_{j,\psi}}\rangle|^2.$$

Hence,

$$\left\| \begin{pmatrix} D_{1,\phi} - D_{1,\psi} + \langle \phi, \Gamma_{\hat{r}_{1,\phi}}\rangle - \langle \psi, \Gamma_{\hat{r}_{1,\psi}}\rangle \\ \vdots \\ D_{s,\phi} - D_{s,\psi} + \langle \phi, \Gamma_{\hat{r}_{s,\phi}}\rangle - \langle \psi, \Gamma_{\hat{r}_{s,\psi}}\rangle \end{pmatrix} \right\|_2 \leq \sum_{j=1}^{s} |D_{j,\phi} - D_{j,\psi}| + |\langle \phi, \Gamma_{\hat{r}_{j,\phi}}\rangle - \langle \psi, \Gamma_{\hat{r}_{j,\psi}}\rangle|$$

since for positive quanities x_i, $\sqrt{\sum_i x_i^2} \leq \sum_i |x_i|$.

From inequality (10)

$$\begin{aligned}&\|G_\phi^{-1}\|_{\text{op}} \left\| \begin{pmatrix} D_{1,\phi} - D_{1,\psi} + \langle \phi, \Gamma_{\hat{r}_{1,\phi}}\rangle - \langle \psi, \Gamma_{\hat{r}_{1,\psi}}\rangle \\ \vdots \\ D_{s,\phi} - D_{s,\psi} + \langle \phi, \Gamma_{\hat{r}_{s,\phi}}\rangle - \langle \psi, \Gamma_{\hat{r}_{s,\psi}}\rangle \end{pmatrix} \right\|_2 \\ &\leq \frac{2\mu}{C_2} \frac{q_{X,T,\phi}^2}{\exp(-M_2^2/(q_{X,T,\phi}^2\mu))T^2} \cdot \\ &s\left[M\|\phi - \psi\|_H + \|\phi\|_H\sqrt{C}T\|\phi - \psi\|_H^{1/2} + \max\|\Gamma_{j,\psi}\|_H\|\phi - \psi\|_H\right], \end{aligned} \tag{12}$$

by an application of Corollary 1 and Theorem 4.5. Here we have used $q_{X,T,\phi}$ to denote the separation distance for the ϕ-related trajectories.

Turning our attention to inequality (11),

$$\begin{aligned}
&\|G_\phi^{-1} - G_\psi^{-1}\|_{\text{op}} \left\| \begin{pmatrix} D_{1,\psi} + \langle \psi, \Gamma_{\hat{r}_{1,\psi}} \rangle \\ \vdots \\ D_{s,\psi} + \langle \psi, \Gamma_{\hat{r}_{s,\psi}} \rangle \end{pmatrix} \right\|_2 \\
\leq &\|G_\phi\|_{\text{op}}^{-1} \|G_\phi - G_\psi\|_{\text{op}} \|G_\psi\|_{\text{op}}^{-1} \cdot \left\| \begin{pmatrix} D_{1,\psi} + \langle \psi, \Gamma_{\hat{r}_{1,\psi}} \rangle \\ \vdots \\ D_{s,\psi} + \langle \psi, \Gamma_{\hat{r}_{s,\psi}} \rangle \end{pmatrix} \right\|_2 \\
\leq &\frac{4\mu^2}{C_2^2} \frac{q_{X,T,\phi}^2 q_{X,T,\psi}^2}{\exp(-((M_2^2/q_{X,T,\phi}^2\mu) + (M_2^2/q_{X,T,\psi}^2\mu))T^4} \cdot \\
&sCT^2 \|\phi - \psi\|_H \left\| \begin{pmatrix} D_{1,\psi} + \langle \psi, \Gamma_{\hat{r}_{1,\psi}} \rangle \\ \vdots \\ D_{s,\psi} + \langle \psi, \Gamma_{\hat{r}_{s,\psi}} \rangle \end{pmatrix} \right\|_2 . \qquad (13)
\end{aligned}$$

Sufficient conditions for convergence: In order to prove convergence of our algorithm we will make some assumptions. Let $1 > \varepsilon > 0$, we will first assume that our iterative procedure starts reasonably close to the true solution F. In terms of the above, we are going to make the assumption that for both ϕ and ψ we have $\|\phi - F\|_H \leq \varepsilon/2$, $\|\psi - F\|_H \leq \varepsilon/2$ and $\|\phi - \psi\|_H \leq \varepsilon < 1$. This in turn gives us that $|D_{j,\phi}|, |D_{j,\psi}| < M\varepsilon$. We will also make an assumption on our underlying vector field. We are going to assume that the vector field F has at-least the smoothness conditions stated in section 3 and is not overly powerful in comparison to the mobile sensor engines. In terms of the above, given the differential equation

$$\dot{p} = s_j \begin{pmatrix} \cos(\theta_j) & \sin(\theta_j) \end{pmatrix}^\top + F(p) \quad p(0) = p_j$$

with solution $r_F(t)$, we will assume $\|\Gamma_{r_F} - \Gamma_\Upsilon\|_H < \varepsilon$ where $\Upsilon(t)$ is the straight line solution to the differential equation above without the F-term, and that $\|\Gamma_\Upsilon\|$ is bounded by some constant independent of T. The latter assumption is not unreasonable, as $\|\Gamma_\Upsilon\|$ would be bounded by the norm of an occupation kernel for the straight line trajectory spanning the diameter of our feature space. With these assumptions, we have $\|\Gamma_{j,\phi}\|_H, \|\Gamma_{j,\psi}\|_H \leq \|\Gamma_\Upsilon\|_H + \varepsilon$. Moreover, by Cauchy Schwarz on $|\langle \phi, \Gamma_{\hat{r}_{j,\phi}} \rangle|$ we have $\left\| \left(D_{1,\phi} + \langle \phi, \Gamma_{\hat{r}_{1,\phi}} \rangle, \cdots, D_{s,\psi} + \langle \phi, \Gamma_{\hat{r}_{s,\phi}} \rangle \right)^\top \right\|_2$ is bounded by some constant L. This above discussion culminates in the following assumptions and lemma.

Assumptions 1. Let F be the true vector field. We make the following assumptions:

1) The functions ϕ and ψ as they appear in the above sections satisfy $\|\phi - F\|_H \leq \varepsilon/2$ and $\|\psi - F\|_H \leq \varepsilon/2$ for some $1 > \varepsilon > 0$
2) If $r_F(t)$ is a trajectory given by the true vector field F, then $\|\Gamma_{r_F} - \Gamma_\Upsilon\| < \varepsilon$ where Γ_Υ is the occupation kernel for the straight line trajectory $\Upsilon(t)$ starting at the same point as $r_F(t)$.
3) For all straight line trajectories, $\Upsilon(t)$, the norm $\|\Gamma_\Upsilon\|$ is bounded by some constant independent of T.

Lemma 4.7. *With the above assumptions,*

$$\left\| \begin{pmatrix} \hat{w}_{1,\phi} \\ \vdots \\ \hat{w}_{s,\phi} \end{pmatrix} \right\|_2 \leq \|G_\phi^{-1}\|_{op} \left\| \begin{pmatrix} D_{1,\phi} + \langle \phi, \Gamma_{\hat{r}_{1,\phi}} \rangle \\ \vdots \\ D_{s,\phi} + \langle \phi, \Gamma_{\hat{r}_{s,\phi}} \rangle \end{pmatrix} \right\|_2 \leq L \cdot \frac{2\mu}{C_2} \frac{q_{X,T,\phi}^2}{\exp(-M_2^2/(q_{X,T,\phi}^2 \mu))T^2}$$

where L is independent of T.

Assumptions 2. Our final assumption is that

$$\|\Gamma_{r_\phi} - \Gamma_{r_\psi}\|_H \leq CT\|\phi - \psi\|_H$$

for some constant C independent of T.

This is an assumption on the regularity on the solutions to the differential equations stated in terms of the occupation kernels. While more work is needed, we postulate that this condition could be stated in terms of Frechet differentiability.

Proof of Theorem 4.1:

Proof. Consider,

$$\begin{aligned} \|A(\phi) - A(\psi)\|_H &= \left\| \sum_{j=1}^{s} \hat{w}_{j,\phi} \Gamma_{\hat{r}_{j,\phi}}(x) - \sum_{j=1}^{s} \hat{w}_{j,\psi} \Gamma_{\hat{r}_{j,\psi}}(x) \right\|_H \\ &\leq \sum_{j=1}^{s} |\hat{w}_{j,\phi}| \|\Gamma_{\hat{r}_{j,\phi}} - \Gamma_{\hat{r}_{j,\psi}}\|_H + \sum_{j=1}^{s} |\hat{w}_{j,\phi} - \hat{w}_{j,\psi}| \|\Gamma_{\hat{r}_{j,\psi}}\|_H \end{aligned}$$

By our assumptions and Lemma 4.7 we have that the first term is proportional to $\frac{1}{T}\|\phi - \psi\|_H$. By inequalities (10), (11) and a reevaluation of our estimates (12) and (13) given the assumptions, the second term in the above is also proportional to $\frac{1}{T}\|\phi - \psi\|_H$. Thus for a suitable T it is a contraction mapping and by Contraction Mapping Theorem convergence of Algorithm 1 is established. □

Remark 6. The inequalities above are also dependent on the number of trajectories and the path separation distance. Specifically, inspection of the above inequalities shows that they are propositional to $sq_{X,T}^2$. However, the main point of the theorem is that assuming those quantities are fixed and/or bounded then if enough data are collected, i.e. for a sufficient amount of time the algorithm is guaranteed to converge. It should be noted that this is a sufficient condition and other conditions may exist in which the algorithm converges.

Remark 7. By default, occupation kernels are universal if the kernels of the RKHS are universal. If $\gamma(t) = y$ for all $t \in [0, T]$ we note that $\Gamma_\gamma(x) = TK(x, y)$. Hence, the set of occupation kernels contains the set of kernels.

5. **Numerical Experiments.** In Experiment 1, we generated a flow field $F(x)$ using a linear combination of Gaussian kernels and generated a set of random points and angles to serve as our anticipated dynamics using unit speed. The flow field is given by

$$F(x) = \frac{1}{8} \begin{pmatrix} f_1(x) \\ f_2(x) \end{pmatrix}, \tag{14}$$

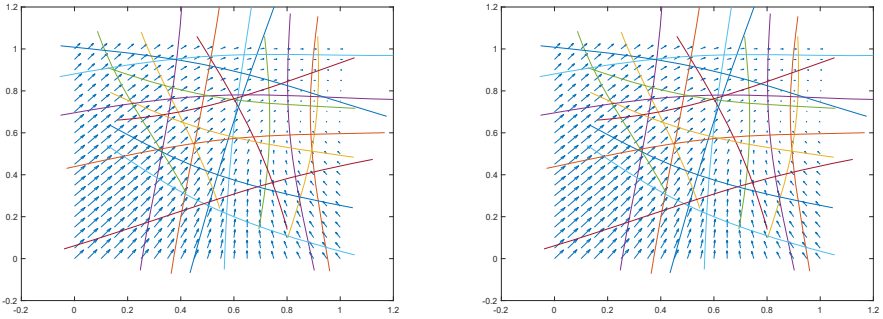
where

$$f_1(x) = 5\exp\left(-2\left\|x - \begin{bmatrix}.25\\.25\end{bmatrix}\right\|^2\right) - .2\exp\left(-\left\|x - \begin{bmatrix}.25\\.75\end{bmatrix}\right\|^2\right) + 2\exp\left(-\left\|x - \begin{bmatrix}.75\\.75\end{bmatrix}\right\|^2\right) - 5\exp\left(-2\left\|x - \begin{bmatrix}.75\\.25\end{bmatrix}\right\|^2\right),$$

and

$$f_2(x) = 3\exp\left(-\left\|x - \begin{bmatrix}.25\\.25\end{bmatrix}\right\|^2\right) + \exp\left(-\left\|x - \begin{bmatrix}.25\\.75\end{bmatrix}\right\|^2\right) - 3\exp\left(-3\left\|x - \begin{bmatrix}.75\\.75\end{bmatrix}\right\|^2\right) + \exp\left(-\left\|x - \begin{bmatrix}.75\\.25\end{bmatrix}\right\|^2\right).$$

The results of Experiment 1 are summarized in Figure 1.

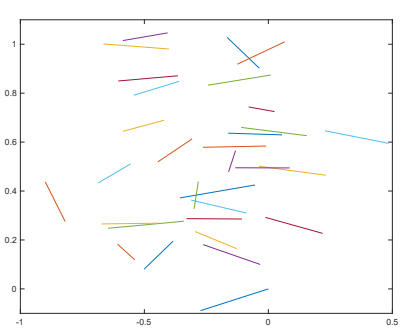


(A) True vector field and trajectories. (B) Results of Experiment 1.

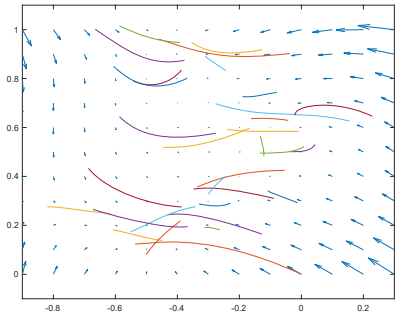
FIGURE 1. Algorithm 1, used for estimation of the vector field in (14). The true trajectories are calculated via RK4 over the time frame $[0, 1]$. Using Gaussian RBFs with a kernel width of 1, we performed 10 iterations of Algorithm 1.

For Experiment 2, summarized in Figures 2a,2b,2c, and 2d, we used the Gliderpalooza 2013 data first presented in [5]. The data were for 31 sequential trajectory segments and contained the initial positions, the true final positions, dead-reckoned final positions, speeds, and dead-reckoning times. For the experiments, the authors of [5] used averaged speeds and an average dead-reckoning time of 3.5 hours.

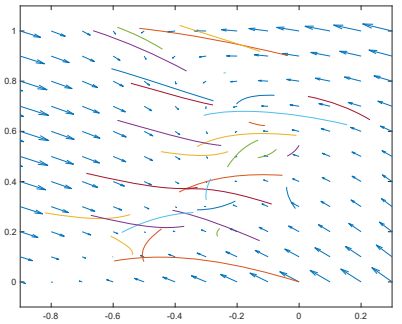
Algorithm 1 is designed to work on multiple trajectories necessitating some prior calculations. The data were scaled by a factor of 10^{-4}, then the provided initial positions and dead-reckoned final positions were used to calculate the initial directions along with calculated speeds using their averaged dead-reckoning time of 3.5 hours. The sequential trajectory data were broken in to 31 separate trajectories. Using exponential kernels, we performed 5, 10, and 20 iterations of Algorithm 1 using a kernel width of $\mu = 1/170$ for the 5 iteration run and a kernel width of $\mu = 1/10,000$ for the 10 and 20 iteration runs. The kernel widths for each of these runs were selected to produce well conditioned Gram matrices.



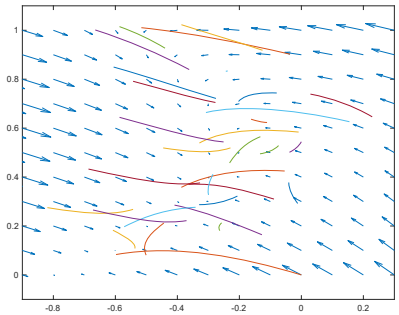
(A) Initial trajectories.



(B) Approximated field and trajectories after 5 iterations.



(C) Approximate field and trajectories with 10 iterations

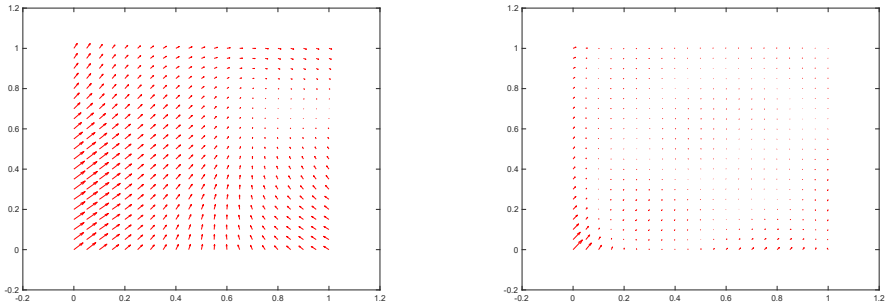


(D) Approximate field and trajectories with 20 iterations

FIGURE 2. Calculated initial trajectories and output of Algorithm 1 for 5, 10, and 20 iterations on Gliderpalooza data.

6. **Discussion.** This approach to motion tomography has several advantages over that of [5]. In the context of this aim, the flow field is approximated using the occupation kernels as basis functions for approximation, whereas [5] requires a piecewise constant description of the flow field or a parameterization with respect to Gaussian RBFs. Moreover, [5] renormalizes the integral by multiplying and dividing by $\|\dot{\tilde{r}}(t)\|_2$ to artificially convert the integrals to line integrals. The renormalization may ultimately produce divide by zero errors that lead to stronger assumptions on the dynamical systems. The occupation kernel method avoids the renormalization and does not add further restrictions on the dynamics. Finally, the representation of the flow field with respect to the occupation kernel basis allows for the application of the approximation powers of RKHSs. It should also be noted that although the applications of this technique apply mainly to $\mathbb{R}^2$ valued functions, there is no inherent reason to limit to $\mathbb{R}^2$. That is, this technique extends to $\mathbb{R}^d$ valued functions.

Figure 3 compares the output of Algorithm 1 with the method in [5] for estimation of the flow field in Experiment 1. Figures 3a, 3b are the plots of the difference



(A) Difference between the true and approximate field given by the method in [5]

(B) Difference between the true and approximate field given by Algorithm 1.

FIGURE 3. A comparison of Algorithm 1 and the method of [5] for estimation of the vector field in (14).

between the true vector field and the approximated vector fields for both methods.

Due to its structure for spatial representation of the flow-field, the method in [5] leads to poor performance for the region where sufficient trajectory information is not available. Quantitatively, given a sample of vectors $\{V(x_1, y_1), \dots, V(x_n, y_n)\}$ and $\{W(x_1, y_1), \dots, W(x_n, y_n)\}$ from two vector fields $V(x, y)$ and $W(x, y)$, we can define the max error (relative to V) as

$$\text{Max Error} = \max_i \{\|V(x_i, y_i) - W(x_i, y_i)\|/\|V(x_i, y_i)\|\} \tag{15}$$

and the mean error (relative to V) as

$$\text{Mean Error} = \text{Mean}\{\|V(x_i, y_i) - W(x_i, y_i)\|/\|V(x_i, y_i)\|\}. \tag{16}$$

Let $V(x, y)$ represent the true vector field and let $W(x, y)$ denote the estimated vector field. Given the samples shown above, the max and mean errors for both methods are summarized in Table 1.

	Method in [5]	Algorithm 1
Max Error	1.1849	0.25321
Mean Error	0.51549	0.025642

TABLE 1. Algorithm 1, along with the technique in [5] are used to estimate the vector field in (14). The table shows the maximum and the average estimation error, as defined in (15) and (16), respectively, for the two methods.

Figure 4 shows the the Mean Error as a function of the number of iterations for three different vector fields. The first vector field is the one used in the first experiment, the second is a linear vector field ($f_1(x) = x_2, f_2(x) = -0.2x_1$) that corresponds to a flow that spirals inwards towards the origin, and the third is a constant vector field ($f_1(x) = 0.2, f_2(x) = 0.1$). None of the three vector fields are in the span of the occupation kernels. The results in Figure 4 demonstrate that the

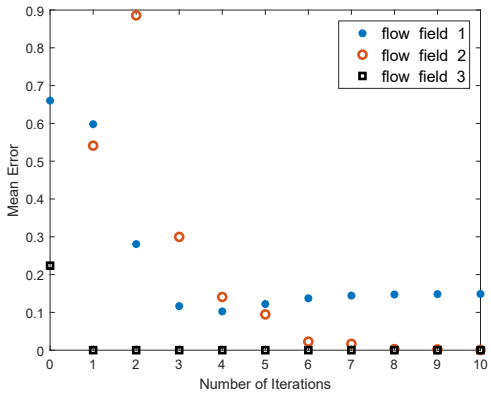


FIGURE 4. To demonstrate universality of the occupation kernel basis, Algorithm 1 is used to estimate three different flow fields. The flow field in (14) (Flow field 1), a linear flow field (Flow field 2), and a constant flow field (Flow field 3). The plot shows the average estimation error, as defined in 16, plotted against the number of iterates of Algorithm 1 for the three flow fields.

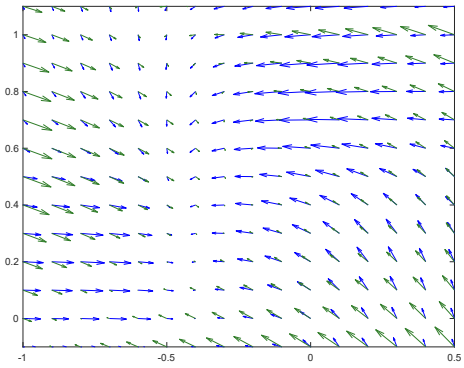


FIGURE 5. Simultaneous plots for estimation of the unknown flow field in the Gliderpalooza experiment using Algorithm 1 (green arrows) and the method of [5] (blue arrows).

developed algorithm can estimate a variety of flow fields, not necessarily in the span of the occupation kernels. In the absence of ground truth in Experiment 2 we show a simultaneous plot of approximated fields from both Algorithm 1 and the method in [5] in Figure 5. For the set of samples depicted in Figure 5, we calculated the maximum norm difference, the mean norm difference, and the variance of the norm differences between the two fields. The results of this comparison are summarized in Table 2.

max norm difference	0.14877
mean norm difference	0.0088628
variance	0.0001869

TABLE 2. Norm difference statistics between estimates of the Gliderpalooza flow field, approximated using Algorithm 1 and the method in [5].

REFERENCES

[1] N. Aronszajn, Theory of Reproducing Kernels, *Transactions of the American mathematical society*, **68** (1950), 337–404.

[2] A. S. Aweiss, B. D. Owens, J. Rios, J. R. Homola and C. P. Mohlenbrink, Unmanned aircraft systems (UAS) traffic management (UTM) national campaign II, in *AIAA Inf. Syst.-AIAA Infotech@ Aerosp.*, 2018, 1727.

[3] Y. Benyamini and J. Lindenstrauss, *Geometric Nonlinear Functional Analysis. Vol. 1*, vol. 48 of American Mathematical Society Colloquium Publications, American Mathematical Society, Providence, RI, 2000.

[4] R. Burden, R. Burden and J. Faires, *Numerical Analysis*, no. v. 1 in Numerical Analysis, Brooks/Cole, 2001.

[5] D. Chang, W. Wu, C. R. Edwards and F. Zhang, Motion tomography: Mapping flow fields using autonomous underwater vehicles, *The International Journal of Robotics Research*, **36** (2017), 320–336.

[6] R. E. Moore, *Computational Functional Analysis*, Ellis Horwood Series: Mathematics and its Applications, Ellis Horwood Ltd., Chichester; Halsted Press [John Wiley & Sons, Inc.], New York, 1985.

[7] V. I. Paulsen and M. Raghupathi, *An Introduction to the Theory of Reproducing Kernel Hilbert Spaces*, vol. 152 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, 2016.

[8] J. Petrich, C. A. Woolsey and D. J. Stilwell, Planar flow model identification for improved navigation of small AUVs, *Ocean Eng.*, **36** (2009), 119–131.

[9] J. A. Rosenfeld, R. Kamalapurkar, L. Gruss and T. T. Johnson, Dynamic Mode Decomposition for Continuous Time Systems with the Liouville Operator, *arXiv preprint arXiv:1910.03977*.

[10] J. A. Rosenfeld, B. Russo, R. Kamalapurkar and T. T. Johnson, The Occupation Kernel Method for Nonlinear System Identification, *arXiv preprint arXiv:1909.11792*.

[11] V. Stepanyan and K. S. Krishnakumar, Estimation, navigation and control of multi-rotor drones in an urban wind field, in *AIAA Inf. Syst.-AIAA Infotech @ Aerosp.*, 2017, 0670.

[12] R. Thompson and W. Walter, *Ordinary Differential Equations*, Graduate Texts in Mathematics, Springer New York, 2013.

[13] H. Wendland, *Scattered Data Approximation*, vol. 17, Cambridge University Press, 2004.

[14] W. Wu, D. Chang and F. Zhang, Glider CT: Reconstructing flow fields from predicted motion of underwater gliders, in *Proceedings of the Eighth ACM International Conference on Underwater Networks and Systems - WUWNet '13,* Article No. 47, 2013, 1–8.

Received xxxx 20xx; revised xxxx 20xx.

E-mail address: russobp@ornl.gov
E-mail address: rushikesh.kamalapurkar@okstate.edu
E-mail address: changdo@oregonstate.edu
E-mail address: rosenfeldj@usf.edu