1 The Ergodicity Metric

In the following discussion, we introduce the concept of the ergodicity metric. Consider a rectangular region $U \subset \mathbb{R}^d$ and assume that the trajectories of the agents are confined within this domain. In [MM21], George Mathew and Igor Mezić define a dynamical system to be ergodic with respect to a measure μ if, for any Borel-measurable subset $U' \subset U$ and any time interval $[0, t]$, the total time a trajectory spends in U' is equal to $\mu(U')$. The measure μ can be defined arbitrarily, depending on the desired sampling or exploration strategy over the domain. However, for this definition to be well-posed, μ must be t-bounded—that is, $0 \leq \mu(U') \leq t$ for all measurable subsets U'.

For the sake of intuition, we derive a related definition in the case where μ is a probability measure over the domain U. In this setting, the system is said to be ergodic if, for any measurable subset $U' \subset U$, the probability that an agent is located within U' at a randomly selected time is equal to $\mu(U')$. Formally, consider a system with a single agent following a trajectory $g : [0, t] \to U$. Then the system is ergodic with respect to the target probability measure μ if and only if, for $\tau \sim \mathcal{U}[0, t]$ and any measurable subset $U' \subset U$, we have:

$$\mathbb{P}(g(\tau) \in U') = \mu(U'). \tag{1.1}$$

For multiple agents with trajectories $g_i, i = 1, 2, \ldots, M$ respectively, this readily extends to

$$\frac{1}{M} \sum_{i=1}^{M} \mathbb{P}(g_i(\tau) \in U') = \mu(U'), U' \subset U \text{ measurable.} \tag{1.2}$$

Now, more formally, we consider the measure spaces $\{U, \Sigma_U, \lambda_g\}$ and $\{[0, t], \Sigma, \lambda\}$, where λ denotes the Lebesgue measure. Let $g : [0, t] \to U$ be a measurable function (cf. Propostion A.2), then in accordance with Definition A.4 we define the measure induced by the trajectory g on U as $\lambda_g := g(\lambda)$. We now state, that by definition the relation:

$$\lambda_g(U') = \lambda(g^{-1}(U')) = \lambda(\tau \in [0, t] \mid g(\tau) \in U'), U' \in \Sigma_U, \tag{1.3}$$

holds. Intuitively, Eq. (1.3) quantifies the amount of time the trajectory g spends in a measurable subset $U' \subset U$. We note that the measure λ_g is not a probability measure, as it is only t-bounded. However, normalizing by the factor $\lambda([0, t])^{-1}$ yields a probability measure. Before we proceed to define the ergodicity metric more formally, we first introduce a technical lemma.

Lemma 1.1. *Let $\{\mu_i\}_{i \in \mathbb{N}}$ be a collection of finite measures, where each μ_i is defined on a corresponding σ-algebra $\mathcal{A}_i$ for each $i \in \mathbb{N}$. Then $\mu = \sum_i \mu_i$ is a measure on the σ-algebra $\mathcal{A} = \bigcap_i \mathcal{A}_i$. The measure space $(S, \mathcal{A}, \mu)$ enjoys the property that for any measurable function $f : S \to \overline{\mathbb{R}}$, the relation:*

$$\int f d\mu = \sum_i \int f d\mu_i \tag{1.4}$$

holds. The symbol $\overline{\mathbb{R}}$ denotes the closure of $\mathbb{R}$. For convenience, and with some abuse of notation, we write $\overline{\mathbb{R}} = [-\infty, \infty]$. Conventions with regard to the symbol ∞ are discussed in Section B.2.

Proof. See Section B.1 □

With the necessary preliminaries in place, we are now prepared to formally define the ergodic property.

Definition 1.1. *Let $\{U, \Sigma_U, \mu\}$ be a probability space and $\{[0, t], \Sigma, \lambda\}$ be measure space, then a dynamical system induced by a set of measurable functions $G := \{g_1, g_2, \ldots g_n\}$, $g_j : [0, t] \to U$, has the ergodic property if and only if*

$$\frac{1}{Nt} \lambda_G = \mu,$$

where λ_G is a measure on U defined under the mappings G, formally $\lambda_G := \sum_{j=1}^{N} g_j(\lambda) := \sum_{j=1}^{N} \lambda_{g_j}$ and λ denotes the Lebesgue measure.

2 Discussion on Ergodic Solutions

In practical applications, the goal is to identify functions G such that the induced dynamical system satisfies the ergodic property as defined in Definition 1.1. As a first step, we consider the natural question: under what conditions can such functions be constructed to ensure ergodicity of the system? To preserve both theoretical clarity and practical applicability, we restrict our analysis to the case where $G := \{g_1, g_2, \dots, g_n\}$ is a collection of continuous functions. For Borel measures and a single continuous trajectory, an answer is provided by the following theorem.

Theorem 2.1 ([Bog07b, Thm 9.7.1. pg. 289])**.** *Let K be a compact metric space that is the image of $[0,1]$ under a continuous mapping f, and let μ be a Borel probability measure on K such that K is its support. Then, there exists a continuous mapping $g : [0,1] \to K$ such that $\mu = \lambda \circ g^{-1}$, where λ is Lebesgue measure on $[0,1]$.*

The following proposition extends the previous Theorem 2.1 to multiple continuous functions G.

Proposition 2.1 (Extension of Theorem 2.1 to multi-agent systems)**.** *Let $\{[0,1], \mathcal{B}([0,1]), \lambda\}$ and $\{K, \mathcal{B}(K), \mu\}$ be measure spaces and let $\{K_i\}_{1 \le i \le n}$ be compact metric spaces that are the image of $[0,1]$ under a continuous mapping f_i, such that $\cup K_i = K$ and $\cap K_i = \emptyset$. Now, let μ be a measure on K, such that there exists Borel probability measures μ_i, which are equivalent to μ on K_i, but vanish everywhere else. Then there exists continuous mappings $G := \{g_1, g_2, \dots g_n\}$, with $g_i : [0,1] \to K_i$ such that $\mu = \sum_{i=1}^n \lambda \circ g_i^{-1}$, where λ is the Lebesgue measure on $[0,1]$. We note that in general $f_i \neq g_i$*

Proof. By the statements of Theorem 2.1, we have that for $\{[0,1], \mathcal{B}([0,1]), \lambda\}$ and $\{K_i, \mathcal{B}(K), \mu\}$, there exists continuous function g such that $\mu_i = \lambda \circ g^{-1}$. Then by Lemma 1.1 and the fact that the K_i's are disjoint, the assertion is shown. □

A discussion on the condition of a metric spaces K_i being the image of continuous function on $[0,1]$ in Proposition 2.1 can be found in Section B.3. We will now give a concrete example of a construction

Example 2.1 (Construction of an trajectory defining an ergodic system)**.** *Let $\{[0,1], \mathcal{B}([0,1]), \lambda\}$ and $\{[0,1]^n, \mathcal{B}([0,1]^n), \lambda^n\}$ be the measure spaces under consideration, where λ^n is the Lebesgue measure on a n-fold Borel-measurable product space. We note that $[0,1]^n$ is the image of a continuous function under a continuous mapping f (cf. Theorem B.1), hence a single trajectory $\{g\}$ suffices to achieve ergodicity. An example of such a mapping is the so called Peano curve, of which the graph is illustrated in Figure 1. In fact it can be shown that this function is measure preserving with respect to the Lebesgue measure that is*

$$\lambda \circ f^{-1} = \lambda^n,$$

where $f : [0,1] \to [0,1]^n$ is the appropriate Peano curve, for a proof we refer to [Mil76, Thm 4.62].

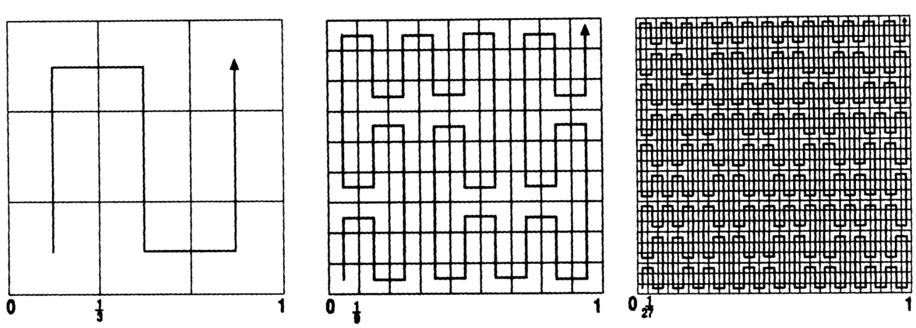


Figure 1: Visualization of the first three elements of a sequence $\{f_k\}_{k \ge 1}$ converging pointwise to the Peano curve. The image was taken from [Sag94, Figure 3.3.2]

Example 2.1 illustrates that even for the simplest Borel measures, a system is ergodic only with respect to highly complex functions. In fact, it is well known that every space-filling curve—such as the Peano curve—has a fractal graph and infinite length. Clearly, no trajectory over a finite time interval can have an infinite-length graph. Therefore, in general, no such solution exists. This observation motivates consideration of other measures, particularly those defined on restricted σ-algebras in higher-dimensional spaces, where suitable solutions might be found. As a motivating example consider the measure space $\{[0,1]^n, \{\emptyset, [0,1]^n\}, \lambda^n\}$, then any trajectory $g : [0,1] \to [0,1]^n$ induces a ergodic system, since in this case a system is ergodic if the trajectory remains within the domain U, which holds by definition. For measures of this type, as in the previous example, the following proposition guarantees the existence of a feasible solution.

Proposition 2.2. *Let $\big([0,1], \mathcal{B}([0,1]), \lambda\big)$ and $\big([0,1]^n, \mathcal{A}, \mu\big)$ be measure spaces where μ is an atomless measure supported on $[0,1]^n$, and $\mathcal{A}$ is an algebra generated by a finite system of closed subsets of $[0,1]^n$. Then there exists a continuous function $g : [0,1] \to [0,1]^n$ such that*

$$\mu = \lambda \circ g^{-1}.$$

Proof. Before stating the proof, we introduce a technical lemma.

Lemma 2.1. *Let $\mathcal{A}$ be an algebra on a set S, generated by a finite collection of closed subsets $\mathcal{E} = \{A_i\}_{i=1}^m$. Then there exists a finite collection of sets $\mathcal{E}' = \{A'_i\}_{i=1}^m$ such that $\mathcal{E}$ and $\mathcal{E}'$ both generate $\mathcal{A}$, and*

$$\bigcup_{i=1}^m A'_i = S, \quad \text{and} \quad A'_i \cap A'_j = \emptyset \quad \text{for } i \neq j.$$

Proof. First, note that adding elements from the algebra $\mathcal{A}$ to a generating system does not change the generated algebra. Formally, if $\mathcal{A}$ is generated by $\mathcal{E} = \{A_i\}_{i=1}^m$, then for any collection $\{B_i\} \subset \mathcal{A}$, the union $\mathcal{E} \cup \{B_i\}$ also generates $\mathcal{A}$. Hence, without loss of generality, we may assume that

$$\bigcup_{i=1} A_i = S.$$

Next, define the collection $\mathcal{E}' = \{A'_i\}_{i=1}$ by

$$A'_i := \overline{A_i \setminus \bigcup_{j \neq i} A_j}.$$

Since $\mathcal{A}$ is an algebra, it follows that each $A'_i \in \mathcal{A}$. Moreover, $\mathcal{E}'$ generates the same algebra $\mathcal{A}$ as $\mathcal{E}$. By construction,

$$\bigcup_{i=1}^m A'_i = S, \quad \text{and} \quad A'_i \cap A'_j = \emptyset \text{ for } i \neq j,$$

which yields the assertion. ☐

As $\mathcal{A}$ was generated by system of subsets of finite cardinality we can conclude that there exists a partition of closed sets $\{A_i\}_{i\geq 1}$ of $[0,1]^n$, which generates $\mathcal{A}$. By a partion of S with closed sets $\{A_i\}_{i\geq 1}$, we mean that $\cup A_i = S$, $\cup_{i:i\neq j} A_i \subset S, j = 1,2,\ldots$ and $\cap int(A_i) = \emptyset$ (cf. Lemma 2.1). Furthermore, as μ has full support we can assume $\mu(A_i) > 0, i = 1,2,\ldots$. Now by convexity of $[0,1]^n$ we state that for any $x, y \in [0,1]^n$, there exist a continuous (linear) function $\gamma : [a,b] \to [0,1]^n$, with $\gamma(a) = x$ and $\gamma(b) = y, [a,b] \subseteq [0,1]$. Now let $\{x_i\}_{i\geq 1}$ be a sequence of points on the boundaries of the A_i's, such that for each set A_i there exists $x_i, x_{i+1} \in A_i \setminus int(A_i)$, with a linear path between the points. Such a sequence can be constructed by choosing an arbitrary sequence $\{x_j\}$ satisfying $x_i, x_{i+1} \in A_i \setminus int(A_i), i = 1,2,\ldots$, then constructing a linear path $\gamma : [0,1] \to [0,1]^n$, with $\gamma(0) = x_i$ and $\gamma(1) = x_{i+1}$, and then joining with the tuple $Y = \{\gamma([0,1]) \cap A_j \setminus int(A_j), j \neq i\}$, that is $\{x_1, \ldots x_i, Y, x_{i+1}, \ldots\}$. We require that Y is ordered such that $\gamma^{-1}(y_k) < \gamma^{-1}(y_l)$ implies $k < l$. This is well defined as γ is injective. Finally, we define the disjoint partition of $[0,1] = \cup I_j$, such that $\sum_{j:x_j,x_{j+1}\in A_i} \lambda(I_j) = \mu(A_i)$, by construction of the sequence $\{x\}_{i\geq 1}$ and the convexity of $[0,1]^n$, we can conclude existence of a function γ, such that $\gamma^{-1}(A_i) = \bigcup_{j:x_j,x_{j+1}\in A_i} I_j$. The assertion then follows by $\cap$- stability of $\{A_i\}_{i\geq 1}$ and the statement of Proposition A.5. ☐

Remark 2.1. *The length of the graph in Proposition 2.2 is upper bounded by $N \max_i d(x_i, x_{i+1})$, where N denotes the number of disjoint subsets generating the algebra described in Lemma 2.1.*

3 Defintion of Ergodic Loss functions

In general it is not clear how to compare measures, as simply checking $\lambda_G(A) = \mu(A)$ for all measurable sets A is exhaustive. To formulate the approach taken in [MM21], we first introduce the notion of a measure-determining set of functions.

Definition 3.1. *Let $\{S, \mathcal{A}\}$ and $\{S', \mathcal{A}'\}$ be measurable spaces. We say a set of $\mathcal{A}$- $\mathcal{A}'$- measurable functions $\{f_k\}_{k\geq 1}$, $f_k : S \to S'$ is measure determining with respect to the set of measures $M(S)$ if and only if*

$$\forall_k \int f_k d\mu = \int f_k d\lambda \implies \mu = \lambda,$$

where $\mu, \lambda \in M(S)$.

Some examples of measure-determining functions are:

- For the measurable space $\{S, \{S, \emptyset\}\}$, the constant function

  $$f : S \to \mathbb{R}, \quad x \mapsto 1,$$

  is measure-determining with respect to the set of finite measures on S, since

  $$\int_S d\mu = \int_S d\lambda \implies \mu(S) = \lambda(S).$$

  For the empty set $\emptyset = S \setminus S$, we have

  $$\mu(S) - \mu(S) = \lambda(S) - \lambda(S) = 0.$$

- By extension of the previous example, for the measurable space $\{S, \{A_1, A_2, \ldots, A_n\}\}$, the set of indicator functions $\{f_k\}_{1\leq k\leq n}$ defined by

  $$f_k = I_{A_k}$$

  is measure-determining.

- For the measurable space $\big(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)\big)$, the set of functions

  $$\big\{\exp\big(i\langle t, x\rangle\big)\big\}_{t\in\mathbb{R}^n}$$

  is measure-determining; see Proposition 3.1.

Proposition 3.1 ([Bog07a, Proposition 3.8.6, pg. 197])**.** *If two finite Borel measures have equal Fourier transforms, then they coincide. In particular, two integrable functions with equal Fourier transforms are equal almost everywhere.*

We now state a necessary and sufficient condition for a set of functions being measure determining over all finite measures on a compact set.

Proposition 3.2. *Let $M(X)$ denote the set of all finite Borel-measures (possibly complex) over X. Then $\{f_k\}$ is measure determining if and only if $C := span(\{f_k\}_k)$ is dense in $L^1(X)$.*

Proof. For sufficiency we consider $\lambda = \mu - \nu$, we then need to show that

$$\forall_{f\in C} \int f d\lambda = \int_X f d\mu - \int_X f d\nu = 0 \implies \lambda = 0, \tag{3.5}$$

by the finiteness of the respective measures this integral always exists and λ is finite aswell. Now as C is dense in L^1 there exists a sequence $\{f_n\}_{n\geq 1} \in C$ such that $f_1 \leq f_2 \leq \cdots$ and $\sup_{n\geq 1} I_A$ for any measurable set A. Now by Proposition A.6 we can write

$$\int I_A d\lambda = \lim_{n\to\infty} \int f_n d\lambda = \lim_{n\to\infty} (\int f_n d\mu - \int f_n d\nu) = \int I_A d\mu - \int I_A d\nu = 0,$$

hence μ and ν agree on all measurable sets, showing the implication.
Assume for the sake of contradiction that the set of functions $\{f_k\}$ is measure-determining, but its linear span C, is not dense in L^1. Then, by the Hahn–Banach theorem [Rud91, Thm. 3.5], there exists a non-zero continuous linear functional

$$T \in (L^1(X))^*$$

such that

$$T(f) = 0 \quad \text{for all } f \in \overline{C},$$

where $\overline{C}$ denotes the closure of C. By the Riesz Representation Theorem [Rud87, Thm 6.19], every continuous linear functional on $L^1(X)$ can be represented by a signed (or finite) Borel measure μ on X, so that

$$T(f) = \int_X f \, d\mu \quad \text{for all } f \in L^1(X).$$

Hence,

$$\int f \, d\mu = 0 \quad \text{for all } f \in \overline{C},$$

but since $T \neq 0$, we have $\mu \neq 0$. However, μ agrees for all $f \in \overline{C}$ with the zero measure. Contradiction! □

In accordance with the previous definition, we may now formulate a more convenient criterion for measuring the ergodicity of a dynamical system, stated in the following proposition.

Proposition 3.3. *Let $\{U, \Sigma_U, \mu\}$ be a probability space and $\{[0,t], \Sigma, \lambda\}$ be measure space and. Furthermore, let $\{f_k\}_{k\geq 1}$ be a set of measure determining functions as per Definition 3.1, then the dynamical system induced by a set of measurable functions $G := \{g_1, g_2, \dots g_n\}$, $g_j : [0,t] \to U$ has the ergodic property if and only if*

$$\forall_k \int f_k d\mu = \frac{1}{tN} \sum_{j=1}^{N} \int f_k \circ g_j d\lambda \tag{3.6}$$

holds. Here λ denotes the Lebesgue measure.

Proof. Before beginning the proof we state a technical Lemma.

Lemma 3.1 ([Bog07a, Thm 3.6.1])**.** *Let (S, A, λ),(S', A', λ') be measure spaces and λ denote the Lebesgue measure. Let g be measurable and define $\lambda' = g(\lambda)$ (cf. Defintion A.4 and Eq. (1.3)). Then a function f is integrable with respect to λ' if and only if $\int f \circ g d\lambda < \infty$, i.e the composition of f and g is integrable. In which case*

$$\int f d\lambda' = \int f \circ g d\lambda$$

holds.

Formally, for the measure defined in Definition 1.1 we obtain the linear evaluation functionals from the proposition in the form:

$$\forall_k \int f_k d\mu = \frac{1}{tN} \int f_k d\lambda_G, \tag{3.7}$$

where λ_G is a measure on U defined under the mappings G, formally $\lambda_G := \sum_{j=1}^N g_j(\lambda) := \sum_{j=1}^N \lambda_{g_j}$. Then by an simple application of the Lemmata 1.1 and 3.1 to the right hand side of Eq. (3.7) we obtain:

$$\begin{aligned} \int f_k d\lambda_G &= \sum_{j=1}^{n} \int f_k d\lambda_{g_j} \text{ by Lemma 1.1} \\ &= \sum_{j=1}^{N} \int f_k \circ g_j d\lambda \text{ by Lemma 3.1.} \end{aligned} \tag{3.8}$$

Finally, the assertion follows by the measure determining property of functions $\{f_k\}$. □

For a discussion of the connection between this framework and the approach presented in [MM21], we refer the reader to Section B.4. As established in Section 2, it is generally not possible to construct a feasible ergodic solution when working under a Borel σ-algebra. Nevertheless, we can define a metric that quantifies the deviation of a given trajectory from ergodicity. This metric takes the form

$$\sum_k l\left(\langle f_k, \mu\rangle, \langle f_k, \lambda_G\rangle\right),$$

where λ_G denotes the measure on U induced by the trajectory mappings G, μ is the target measure, and l is an arbitrary loss function.

We emphasize that various structurally distinct loss functions have been successfully employed in the literature. For instance, the authors of [SGTM25] employed the inner product loss $\langle \mu-\lambda_G, \mu-\lambda_G\rangle$.

4 Optimization

We now restate the key results from [MM21] concerning the optimization of the ergodic metric expressed in its measure-determining functional form.

Consider the cost function defined as

$$\Phi(t) = \frac{1}{2}\sum_k \left(\langle f_k, \lambda_G\rangle - |G|t\,\langle f_k, \mu\rangle\right)^2 = \frac{1}{2}\sum_k |S_k(t)|^2,$$

where $S_k(t) := \langle f_k, \lambda_G\rangle - |G|t\langle f_k, \mu\rangle$ and μ denotes the target measure. To facilitate optimization, we perform a second-order Taylor expansion of Φ around time t, yielding

$$\Phi(t+\Delta t) = \Phi(t) + \dot\Phi(t)\Delta t + \frac{1}{2}\ddot\Phi(t)(\Delta t)^2 + \mathcal{O}((\Delta t)^3).$$

The first and second derivatives of Φ can be expressed as

$$\dot\Phi(t) = \sum_k S_k(t)\dot S_k(t),$$

$$\ddot\Phi(t) = \sum_k \dot S_k(t)^2 + \sum_k S_k(t)\ddot S_k(t).$$

The derivatives of S_k with respect to τ are given by:

$$\dot S_k = \sum_{g\in G}\langle f_k, \lambda_{g(\tau)}\rangle - |G|t\langle f_k, \lambda\rangle,$$

$$\text{and } \ddot S_k = \sum_{g\in G}\nabla\langle f_k, \lambda_{g(\tau)}\rangle\, Dg,$$

where $Dg = f(g,u)$ represents the state derivative of the agents g, which depends on the control input u.

Substituting these expressions back into the expansion, we obtain

$$\ddot\Phi(t) = \sum_k \dot S_k(t)^2 + \sum_k S_k(t)\sum_{g\in G}\left\langle \nabla\langle f_k, \lambda_g\rangle, f(g,u)\right\rangle.$$

Since only the second term depends explicitly on the control u, the optimization problem reduces to minimizing

$$\sum_k S_k(t)\sum_{g\in G}\left\langle \nabla\langle f_k, \lambda_{g(\tau)}\rangle, f(g,u)\right\rangle,$$

subject to the constraint

$$\|u_g\| \le u_{\max}.$$

In the case of first-order dynamics, where $Dg = u$, this expression simplifies, and the optimal control can be found by solving

$$\min_u \sum_k S_k(t)\sum_{g\in G}\left\langle \nabla\langle f_k, \lambda_g\rangle, u\right\rangle, \quad \text{with} \quad \|u\| \le u_{\max}.$$

This linear optimization is minimized by selecting the control

$$u_g^* = -u_{\max} \frac{\sum_k \nabla \langle f_k, \lambda_g \rangle}{\|\sum_k \nabla \langle f_k, \lambda_g \rangle\|}.$$

We note that even higher order dynamics can be optimized in this fashion but more effort is required for deriving feedback laws.

5 Examples of measure determining sets of functions $\{f_k\}$

In conjunction with Proposition 2.2 the set of indicator function $\{I_{A_k}\}_{1\leq k\leq N}$, is measure determining over the Algebra $\{A_1, A_2, \ldots, A_N\}$. In accordance with the results derived in Section 4 it is evident that these basis are unfeasible for the task at hand, as their gradient evaluates to zero almost everywhere. However, we may still approximate them by a smooth function. That is let $\{X, d_X\}$ be a metric space and $A \subset X$ be a set on which we wish to approximate the Indicator function we then define $A_\epsilon = \{x \in X \mid \mathrm{dist}(x, A) \leq \epsilon\}$, where $\mathrm{dist}(x, A) = \sup_{y\in A} d_X(x, y)$. We write for the approximation of the indicator function over A_ϵ:

$$\tilde{I}_{A_\epsilon}(x) = \begin{cases} 1 & \text{if } x \in A \\ h(x) & \text{if } x \in A_\epsilon \setminus A \\ 0 & \text{else} \end{cases}, \tag{5.9}$$

where h is chosen such that $\tilde{I}$ is smooth. As an example for $A = [a, b] \subset \{\mathbb{R}, |\cdot|\}$ we get

$$\tilde{I}_{[a-\epsilon,b+\epsilon]}(x) = \begin{cases} 1 & \text{if } x \in [a, b] \\ g(x) & \text{if } x \in [b, b+\epsilon] \\ h(x) & \text{if } x \in [a-\epsilon, a] \\ 0 & \text{else.} \end{cases}$$

Intuitively this definition allows for a smooth transition between the respective sets as the boundaries are smooth. An example of such an approximation can be observed in Figure 2.

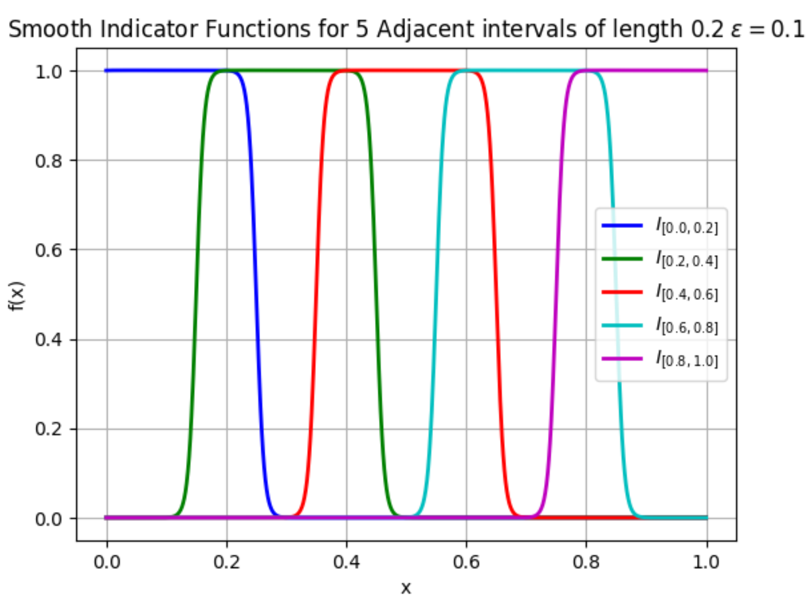


Figure 2: Smooth estimation of indicator functions

One may interpret the smooth approximation of indicator functions as a linear combination of elementary indicator functions that converges uniformly. Specifically, the smooth approximation of the indicator function, as introduced in Eq. (5.9), can be expressed as

$$f_{N,\epsilon} = \sum_{i=0}^{N} c_i I_{A_{\frac{\epsilon i}{N}}},$$

where $\{a_i\}_{i\in\mathbb{N}}$ is a decreasing sequence satisfying $\lim_{i\to\infty} a_i = 0$ and $\sum_i a_i = 1$, with $A_0 = A$. In the limit, we have

$$\lim_{N\to\infty} f_{N,\epsilon} = \tilde{I}_{A_\epsilon},$$

where $f_{N,\epsilon}$ converges pointwise to a smooth function with respect to N. Moreover, taking $\epsilon \to 0$, we recover the original indicator function:

$$\lim_{\epsilon\to 0} f_{N,\epsilon} = I_A.$$

The coefficients a_i can be understood as encoding the relative importance of approximating the indicator function on the subsets $A_{\frac{\epsilon i}{N}}$, with the highest weight placed on the original set A. An illustration of this structure is provided in Figure 3.

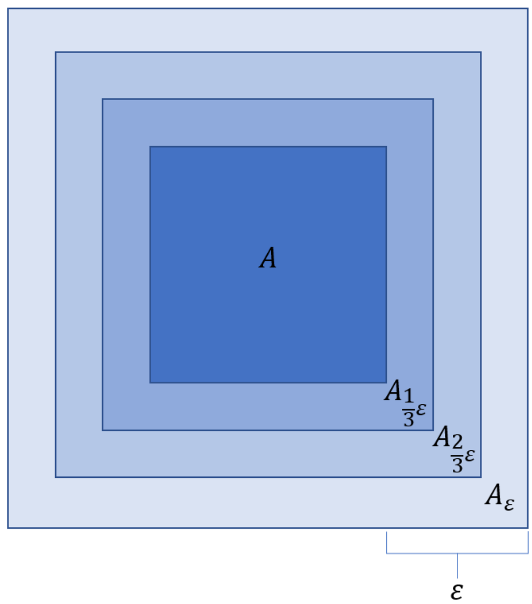


Figure 3: Representation of the sets on which $f_{4,\epsilon}$ takes a unique value. The shade of each set represents the magnitude c_i of the function over this set the brighter the shade the lower the value. The sets are contained within the normed linear space $\{\mathbb{R}^2, \|\cdot\|_\infty\}$

A natural question that arises is whether we can choose sets A, the parameter ϵ and the coefficients $\{a_i\}$ in a manner that yields the measure-determining property over an arbitrary algebra. In the case of the Borel-σ-algebra, we will derive the indicator function of a single point from a family of sets. Specifically, we consider sets centered at a collection of points $\{c_i\}$, defined by

$$A_{\epsilon,c_k} = \{x \in X \mid \|x - c_k\|_p \leq \epsilon\}.$$

Figure 3 illustrates such sets using the L_∞ norm. We now proceed to analyze how these sets can be used to construct the indicator function of a point, for that matter consider:

$$f_{c_k,N,\epsilon} = \sum_{i=0}^{N} a_i I_{A_{\frac{ri}{N}}},$$

then there exists some sequence a_i, $i \in \mathbb{N}$, such that $f_{c_k,N,\epsilon}$ converges point wise to a function of the form $f_{c_k} : \mathbb{R}^n \to \mathbb{R} : x \mapsto \varphi(\|x - c_k\|_p)$, where $\varphi : \mathbb{R}^+ \to \mathbb{R}^+$ is strictly decreasing. The set of functions $\{f_{c_k}\}_{k\in\mathbb{N}}$ are called the radial basis function. In fact it can be shown that these basis form a dense space in the continuous functions over a compact domain for $p \in [1, 2]$ [Bax10]. We can therefore conclude, by Proposition 3.2 together with the fact that the continuous functions form a dense subspace of L_1, that $\{f_{c_k}\}$ is measure determining on the Borel σ-Algebras. In conclusion, we have motivated the use of radial basis functions as approximations of indicator functions centered at single points.

Specifically, we propose the following sets of functions as approximations to the measure-determining property, noting that this is necessarily an approximation since computing an infinite number of coefficients is generally infeasible. These sets include smooth approximations of indicator functions over certain regions, as previously motivated by Proposition 2.2. In particular, when approximating indicator functions of single points, the limiting case yields a family of radial basis functions, which we also propose as a noteworthy and measure-determining set of functions. Finally, drawing on the results of [MM21] and Proposition 3.1, we also include the Fourier basis functions as a further example of a measure-determining family.

6 Simulation

In this section, we evaluate the effectiveness of the proposed measure-determining sets of basis functions (cf. Section 5) using the optimization framework described in Section 4. Specifically, we perform the optimization procedure outlined in Section 4 and compute the coefficients $c_k = \langle f_k, \lambda_G \rangle$ corresponding to the optimized trajectory-induced measure λ_G. These coefficients are then projected onto the respective basis function set, yielding new coefficients $\tilde{c}_k$. From these, we construct an approximate measure $\tilde{\lambda}_G$ defined by

$$\tilde{\lambda}_G(A) = \left\langle I_A, \sum_k \tilde{c}_k f_k \right\rangle .$$

We assess the accuracy of this approximation and the associated optimization performance using several evaluation metrics, detailed below and motivated in part by the framework introduced in [Ope25].

- L^2 **distance** between two probability density functions (pdfs) p and q:

  $$L^2(p, q) = \|p - q\|_2 = \sqrt{\int (p(x) - q(x))^2 \, dx}$$

  Measures the root mean squared difference between the densities. It is sensitive to large pointwise differences but treats all discrepancies uniformly. However, it may not capture perceptual similarity well if differences occur in low-density regions.

- **Total Variation (TV) distance**:

  $$\mathrm{TV}(p, q) = \frac{1}{2} \int |p(x) - q(x)| \, dx$$

  Represents the maximal difference in probability mass assigned to any event. It has a clear probabilistic interpretation and is bounded between 0 and 1, but can be overly sensitive to small local fluctuations.

- **Kullback-Leibler (KL) divergence** from p to q:

  $$\mathrm{KL}(p\|q) = \int p(x) \log \frac{p(x)}{q(x)} \, dx$$

  Measures the information lost when q is used to approximate p. It is asymmetric and heavily penalizes regions where q underestimates p. Useful for inference and likelihood-based methods but undefined if $q = 0$ where $p > 0$.

- **Jensen-Shannon (JS) divergence** (a symmetric and smoothed variant of KL):

  $$m(x) = \frac{p(x) + q(x)}{2}, \quad \mathrm{JS}(p, q) = \frac{1}{2}\mathrm{KL}(p\|m) + \frac{1}{2}\mathrm{KL}(q\|m)$$

  Provides a symmetric and finite divergence measure, smoothing the KL divergences by averaging. It is more robust to zero probabilities and interpretable as a measure of similarity.

- **Cosine similarity**:

$$\mathrm{CosSim}(p, q) = \frac{\int p(x)q(x)\,dx}{\sqrt{\int p(x)^2\,dx}\sqrt{\int q(x)^2\,dx}}$$

Measures the angle between p and q when viewed as vectors in function space, capturing similarity in shape rather than magnitude. It ranges from 0 (orthogonal) to 1 (identical up to scaling) and is insensitive to scale differences.

Figure 4 illustrates the performance of different sets of basis functions with respect to the evaluation metrics introduced previously. As can be seen, both the radial basis functions and the smooth indicator functions converge faster than the Fourier basis. This behavior is expected since the Fourier basis functions are not equally weighted across the domain (cf. [MM21]).

Although the Fourier basis functions exhibit slower initial progress, they continue to improve with extended optimization, beyond what is shown in Figure 4. The radial basis and smooth indicator functions, however, tend to plateau quicker (this behavior is not observed for the Kullback-Leibler divergence performance metric). Therefore, we conclude that, while the Fourier basis ultimately yields the best results given sufficient time, the radial and indicator-based approaches offer superior performance over shorter time horizons. This property is advantageous for time-critical applications, such as search-and-rescue missions involving moving targets or adaptive target measures, where rapid convergence is essential.

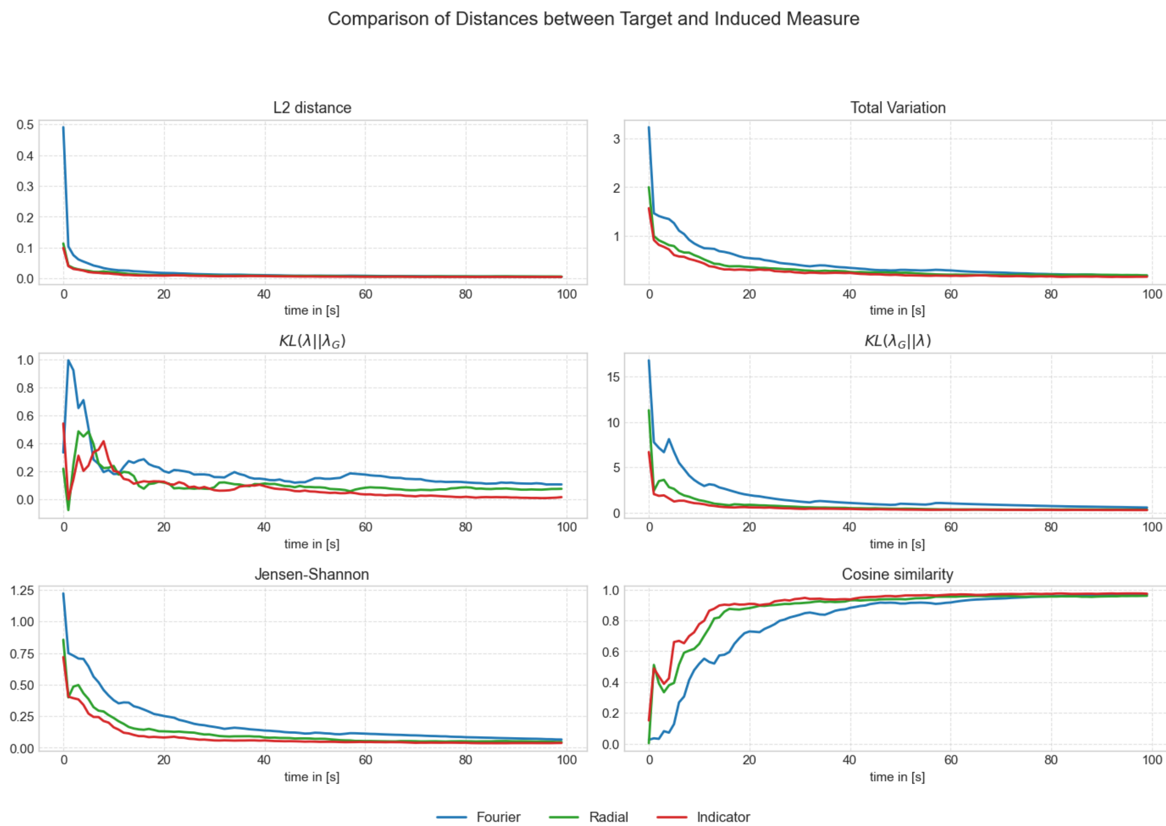


Figure 4: Several evaluation metrics for comparing similarity between probability measures over time on a $[0, 1] \times [0, 1]$. Each set of basis functions consists of 50×50 individual basis. The target measure is described by a gaussian mixture model. The system is governed by first order dynamics with $u_{max} = 0.2$

Another notable advantage of the alternative basis functions introduced in this work is their natural extensibility to non-Euclidean spaces, such as $SE(2)$, $SO(3)$, or more generally $SU(n)$, which frequently arise in control theory. While previous works (([MM13])), have explored the extension of ergodic control to $SE(2)$ using Fourier-based methods, it is acknowledged [SGTM25] that such extensions are generally non-trivial when relying on the Fourier basis. This highlights a potential advantage of the alternative basis functions proposed here, which may extend to these spaces more directly and with fewer technical complications.

We conclude this section by noting that the proof of Proposition 2.2 is constructive. Designing explicit feedback laws—for instance, by connecting the boundaries of the respective partition sets using Dubins paths—presents an interesting direction for future research.

7 Generalization of the ergodic framework

In the following we shall give a more general definition of the ergodicity-property. To this end, we extend the ergodic measure to incorporate sensor information. Formally, we arrive at the following definition.

Definition 7.1 (Measure-valued Trajectories). *Let $\{U, \mathcal{B}(U)\}$ be a measurable space and let $\mu : U \to \mathcal{P}(U)$*[1] *$: x \mapsto \mu_x$ be a weak*-measurable (cf. Definition A.1), measure valued function. Additionally, let $g : [0,t] \to U$, be the continuous trajectory of an agent. The agent is then said to induce the measure $t \mapsto \frac{1}{t}\int_{[0,t]} \mu_{g(\tau)} d\tau := \lambda_g$ at time t.*

From this definition, we may express the generalized ergodic condition for a single-agent system as follows: the system is said to be ergodic if and only if

$$\frac{1}{t}\lambda_g = \nu,$$

where ν denotes the target measure.

It is important to note that if we neglect the assumption of weak*-measurability (cf. Definition 7.1)—by instead requiring measurability with respect to the strong topology—we would inadvertently exclude many relevant and important examples, as outlined in the following remark.

Remark 7.1. *The map $\mu : U \mapsto \mathcal{P}(U)$ needs to be weak* measurable (cf. Definition A.1). As for example the measure valued mapping $t \mapsto \delta_{g(t)} := \delta_x \circ g(t)$ is discontinuous at every point t where g is not locally constant. If the set of such discontinuities has nonzero measure, the mapping fails to be strongly measurable. In contrast, under the weak-* topology, the mapping $t \mapsto \delta_{g(t)}$ is measurable whenever $t \mapsto g(t)$ is measurable.*

Furthermore , in the following two propositions we highlight important properties of the newly defined sensor measure λ_g.

Proposition 7.1. *The measure λ_g, induced by an agent, as per Definition 7.1, indeed defines a Borel-probability measure.*

Proof. Monotonicity and σ-continuity are immediate by the fact that each μ_x is a valid measure. The σ-sub additivity follows for disjoint $\{A_i\}$, per the relation

$$\begin{aligned}
\lambda_g(\bigcup_i A_i) &\leq \sum_i \lambda_g(A_i) \\
\frac{1}{t}\int_{[0,t]} \mu_{g(\tau)}(\bigcup_i A_i) d\tau &\leq \sum_i \frac{1}{t}\int_{[0,t]} \mu_{g(\tau)}(A_i) d\tau \\
\int_{[0,t]} \mu_g(\tau)(\bigcup_i A_i) d\tau &\leq \int_{[0,t]} \sum_i \mu_g(\tau)(A_i) d\tau,
\end{aligned}$$

which yields the assertion since each $\mu_{g(\tau)}$ is a valid measure. The change of summation and integration on the right hand side is justified by Tonellis theorem A.1. We can employ the theorem as $\tau \mapsto \mu_{g(\tau)}(A)$ is non-negative, bounded and measurable on the finite interval $[0,t]$. This follows from the fact that

$$\tau \mapsto \mu_{g(\tau)}(A),$$

[1] $\mathcal{P}(U)$ denotes the set of Borel probability measures over U

is measurable as per the weak*-measurability assumption and

$$x \mapsto \int_A 1\, d\mu_x = \mu_x(A),$$

is measurable. Finally, $g : [0,t] \to U$ is measurable by continuity and the composition of two measurable functions is measurable. $\square$

Proposition 7.2. *Let λ_g be a measure; induced by an agent, as per Definition 7.1, then we have that for any measurable function $f : U \to \mathbb{R}$ with $\int_0^t \int_U |f(x)|\, d\mu_{g(\tau)}(x)\, d\tau < \infty$, the following holds:*

$$\int_U f(x)\, d\lambda_g(x) \;=\; \int_0^t \int_U f(x)\, d\mu_{g(\tau)}(x)\, d\tau.$$

Proof. 1. *Simple functions.* Let $s(x) = \sum_{k=1}^N a_k\, \mathbf{1}_{A_k}(x)$ be a simple function. Then

$$\int_U s\, d\lambda_g = \sum_{k=1}^N a_k\, \lambda_g(A_k) = \sum_{k=1}^N a_k \int_0^t d\mu_{g(\tau)}(A_k)\, d\tau = \int_0^t \sum_{k=1}^N a_k\, d\mu_{g(\tau)}(A_k)\, d\tau = \int_0^t \int_U s\, dd\mu_{g(\tau)}\, d\tau,$$

change of integration and summation follows exactly as in the proof of Proposition 7.1.
2. *Nonnegative measurable f.* Given $f \geq 0$, choose an increasing sequence of simple functions $s_n \uparrow f$. By monotone convergence,

$$\int_U f\, d\lambda_g = \lim_{n\to\infty} \int_U s_n\, d\lambda_g = \lim_{n\to\infty} \int_0^t \int_U s_n\, d\mu_{g(\tau)}\, d\tau = \int_0^t \int_U f\, d\mu_{g(\tau)}\, d\tau,$$

exactly as in the proof of Lemma 1.1.

3. *General integrable f.* If f satisfies $\int_0^t \int_U |f|\, d\mu_{g(\tau)}\, d\tau < \infty$, write $f = f^+ - f^-$ and apply the above to f^+ and f^-. Subtracting gives the desired identity (for more details we refer the reader to the proof of Lemma 1.1). $\square$

With these new definitions in mind we aim to explore an alternative approach to trajectory optimization in ergodic exploration by elaborating on the unicycle example presented in [MM13]. In particular, we consider a kinematic model frequently used to represent mobile robotic agents with non holonomic constraints. The state of the system is defined as

$$g(t) = \begin{bmatrix} X(t) \\ Y(t) \\ \theta(t) \end{bmatrix},$$

where $X(t)$ and $Y(t)$ represent the Cartesian coordinates of the agent in the plane, and $\theta(t)$ denotes the agent's heading angle measured counterclockwise from the global X-axis.

The control input is given by

$$u(t) = \begin{bmatrix} v(t) \\ \omega(t) \end{bmatrix},$$

where $v(t)$ is the translational (forward) velocity and $\omega(t)$ is the angular (rotational) velocity.

The time evolution of the state is governed by the unicycle model:

$$\dot{g}(t) = \begin{bmatrix} \cos(\theta(t)) & 0 \\ \sin(\theta(t)) & 0 \\ 0 & 1 \end{bmatrix} \cdot u(t) := M(t)u(t).$$

Additionally we define the measure valued function the agent induces at position x (cf. Definition 7.1) as:

$$\mu_x : (x, y, \theta) \mapsto \frac{1}{\pi r^2} I_{B(\{x + r\cos(\theta), y + r\sin(\theta)\}, r)} \lambda, \tag{7.10}$$

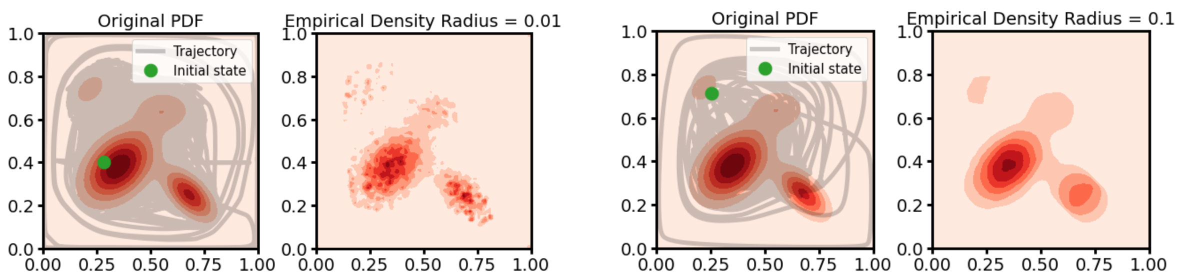
where λ denotes the Lebesgue measure and $B(x, r) = \{y \in U : \|x - y\|_2 \leq r\}$. Then following the steps outlined in Section 4 we obtain

$$\min_u \sum_k S_k(t) \sum_{g \in G} \nabla \langle B_g(t), Dg \rangle, \quad \text{with} \quad \|u\| \leq u_{\max},$$

where $B_g(t) = \langle f_k, \lambda_g \rangle$ This linear optimization is minimized by selecting the control

$$u_g^* = -u_{\max} \frac{\sum_k \langle \nabla B_g(t), M(t) \rangle}{\|\sum_k \langle \nabla B_g(t), M(t) \rangle\|}.$$

Figure 5 shows the results obtained by this feedback law, where the set $\{f_k\}$ was choose to consist of Fourier-basis as in [MM21]. Readers who are concerned with the analytical form of $\langle f_k, \lambda_g \rangle$ are referred to [GS64, pg. 198]



(a) (b)

Figure 5: Optimization over 100000 steps. The empirical distribution was obtained by employing Monte Carlo integration to the measure defined in Definition 7.1

With the previously introduced formulation (cf. Definition 7.1), a key challenge arises: the inner product $\langle f_k, \lambda_g \rangle$ may not be available in analytical form for a general measure-valued function μ_x, unlike the special case $\mu_x = \delta_x$.

In scenarios where, for any x, the support of μ_x is small compared to the support of the target measure ν, i.e., $\text{supp}(\mu_x) \ll \text{supp}(\nu)$, a uniform, geometrically motivated approximation of the underlying density of μ_x appears reasonable. However, when the support of the sensor measure μ_x is significant relative to that of the target, this approximation becomes unreliable. Moreover, numerical computation of such integrals would be computationally expensive and would render the application of the ∇-operator impractical.

Assuming that the mapping μ_x assigns structurally equivalent measures to each point x — for instance, as in Eq. (7.10) — one can exploit standard properties of the Fourier transform. We now introduce the relevant properties of the Fourier transform that enable this approach.

- **Translation (Shift) Theorem:** Translating a function $f(x, y)$ in space by a vector (x_0, y_0) results in a phase shift in the Fourier domain:

$$f(x - x_0, y - y_0) \quad \stackrel{\mathcal{F}}{\longleftrightarrow} \quad \hat{f}(k_x, k_y) \cdot e^{-i2\pi(k_x x_0 + k_y y_0)}.$$

  This means that once the Fourier coefficients of a function centered at the origin are computed, translating the function simply requires multiplying the spectrum by a complex exponential phase factor.

- **Rotation Theorem:** Rotating a function $f(x, y)$ by an angle α corresponds to rotating its Fourier transform by the same angle:

$$f(R_\alpha^{-1}\mathbf{x}) \quad \stackrel{\mathcal{F}}{\longleftrightarrow} \quad \hat{f}(R_\alpha^{-1}\mathbf{k}),$$

  where the rotation matrix is given by

$$R_\alpha = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}.$$

In practice, this means that if a function is rotated in the spatial domain, its frequency content is also rotated accordingly in the Fourier domain.

Together, these properties imply that once the Fourier transform of a base function (e.g., centered and unrotated) is computed, the transform of any translated and rotated version can be obtained analytically by combining phase modulation and frequency-space rotation. The Fourier coefficients of the underlying density rotated by an angle α and translated to the point (x_0, y_0) are then given by the following expression:

$$\hat{f}_{transformed}(\boldsymbol{k}) = \hat{f}(R_{-\alpha}^{-1}\boldsymbol{k})e^{-i2\pi(k_x x_0 + k_y y_0)} \tag{7.11}$$

This avoids recomputing the full transform and enables efficient gradient computations with respect to the translation and rotation parameters. The gradient with respect to $\{x_0, y_0, \alpha\}$ is then given by:

$$\nabla \hat{f}_{\text{transformed}}(k_x, k_y) = e^{-i2\pi(k_x x_0 + k_y y_0)} \cdot \begin{bmatrix} -i2\pi k_x \cdot \hat{f}(R_{-\alpha}\boldsymbol{k}) \\ -i2\pi k_y \cdot \hat{f}(R_{-\alpha}\boldsymbol{k}) \\ \nabla_{\boldsymbol{k}'} \hat{f}(\boldsymbol{k}') \cdot \dfrac{\partial \boldsymbol{k}'}{\partial \alpha} \end{bmatrix},$$

$$\text{where} \quad \boldsymbol{k} = \begin{bmatrix} k_x \\ k_y \end{bmatrix}, \quad \boldsymbol{k}' = R_{-\alpha}\boldsymbol{k}, \quad R_{-\alpha} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}, \quad \frac{\partial \boldsymbol{k}'}{\partial \alpha} = \begin{bmatrix} -k_x \sin\alpha + k_y \cos\alpha \\ -k_x \cos\alpha - k_y \sin\alpha \end{bmatrix}.$$

Accordingly, the results of Section 4 can be utilized. It is imperative to recognize that, in most cases, the function $f(\boldsymbol{k})$ is not available and must be estimated. Before leveraging the method introduced in Section 4, we shall also take into consideration the imaginary parts of the respective Fourier coefficients by incorporating them into the loss function. To this end, we define the new loss function as:

$$\Phi(t) = \frac{1}{2}\sum_k \left(\Re[\langle f_k, \lambda_g\rangle] - \Re[t\,\langle f_k, \lambda\rangle]\right)^2 + \frac{1}{2}\sum_k \left(\Im[\langle f_k, \lambda_g\rangle] - \Im[t\,\langle f_k, \lambda\rangle]\right)^2,$$

where $f_k = e^{-i2\pi(k_x x + k_y y)}$ and λ denotes the target measure. At this juncture, it is possible to leverage the results from Section 4 directly to derive the feedback law:

$$u_g^* = -u_{\max} \frac{\sum_k \nabla\Re[\langle f_k, \lambda_g\rangle] + \nabla\Im[\langle f_k, \lambda_g\rangle]}{\|\sum_k \nabla\Re[\langle f_k, \lambda_g\rangle] + \nabla\Im[\langle f_k, \lambda_g\rangle]\|}.$$

We leave the development of alternative, potentially more suitable formulations of the loss function to future work—for instance, a formulation based on the L_2 distance between the coefficients. The following measure-valued function, characterized by an angle-dependent and non-symmetric density, will be used to evaluate the efficiency of the proposed approach:

$$(x, y, \theta) \mapsto f(x, y, \mu_\theta)\lambda, \tag{7.12}$$

$$\text{where} \quad f(x, y, \mu_\theta) = \frac{1}{2\pi\,\sigma^2\, I_0(\kappa)} \exp\left(-\frac{(x - x_0)^2 + (y - y_0)^2}{2\sigma^2}\right) \cdot \exp\left(\kappa \cos\left(\theta(x, y) - \mu_\theta\right)\right),$$
$$\text{with} \quad \theta(x, y) = \arctan 2(y - y_0, x - x_0)$$

and I_0 denotes the modified Bessel-function of the first kind. The rotation and translation of this function to an arbitrary point in the domain is visualized in Figure 6.

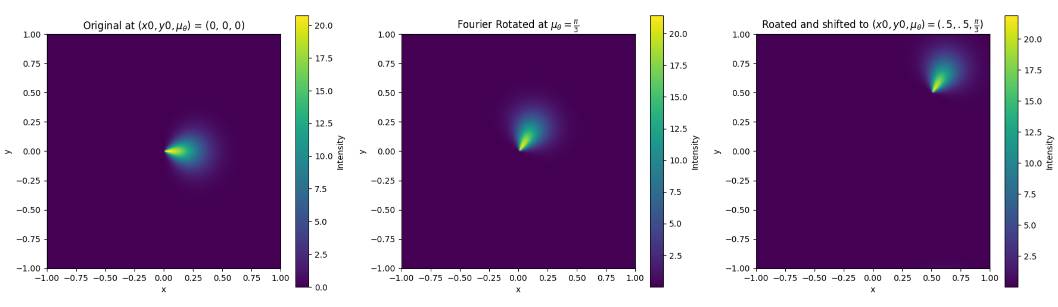


Figure 6: Rotation and translation of the objective function in the Fourier domain

Figure 7 illustrates the trajectory generated by the proposed method. This trajectory demonstrates how the system evolves over time under the derived feedback law, highlighting the effectiveness of our approach in guiding the dynamics towards the target distribution.

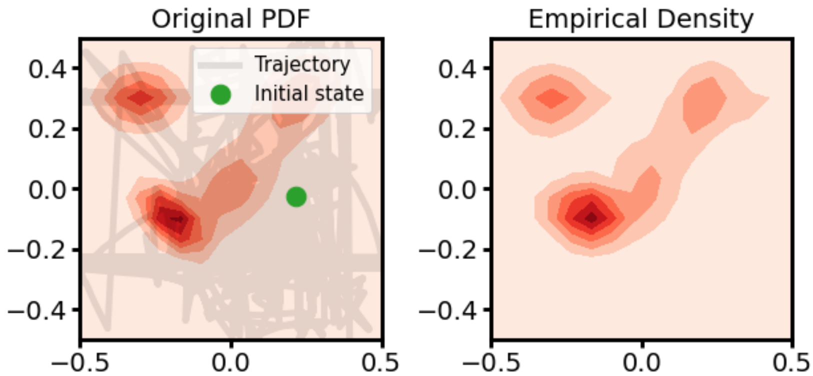


Figure 7: Sampling a Gaussian mixture density using the sensor measure defined in Eq. 7.12 with parameters $\kappa = 5$ and $\sigma = 0.1$. The empirical density was obtained via Monte Carlo integration of the Fourier coefficients followed by the inverse transform. The trajectory was simulated over 200,000 time steps, each corresponding to 0.01 seconds. 16×16 Fourier coefficients were used for evaluation.

8 Conclusion

In this work, we rigorously introduced the concept of measure-determining sets of functions. We also established a necessary and sufficient condition for these sets to determine finite Borel measures. In this context, we examined three notable classes of functions: Fourier basis functions, radial basis functions, and smooth approximations of indicator functions. To evaluate their practical applications, we analyzed how the ergodic metric converges when using these sets. Our findings suggest that Fourier basis functions exhibit superior asymptotic convergence characteristics, whereas radial and smoothed indicator functions converge more quickly initially but saturate sooner. This behavior suggests their particular suitability for time-sensitive applications, such as the pursuit of dynamically evolving target measures, where rapid exploratory behavior is prioritized over long-term thoroughness. In the second part of this paper, we extended the previously defined ergodic metric to incorporate sensor-based information in a mathematically rigorous manner. Building on this extension, we generalized the associated feedback control laws to accommodate sensor-dependent formulations. To address the challenges posed by highly nontrivial sensor measurements, we proposed a principled, computationally efficient solution. We demonstrated the effectiveness of this approach through a representative example, highlighting its practical applicability in complex sensing environments.

A Auxiliary results

A.1 Measurability

Proposition A.1 (Generated σ-algebras [BK13, pg. 8])**.** *For every system $\mathcal{E}$ of subsets in S, there exists a smallest σ-algebra which contains $\mathcal{E}$. We obtain it as the intersection of all σ-algebras containing $\mathcal{E}$:*

$$\sigma(\mathcal{E}) = \{A \subset S \mid A \subset \tilde{A} \text{ for every } \sigma\text{-algebra } \tilde{A} \text{ in } S \text{ with } \tilde{A} \supset \mathcal{E}\ \}$$

Proposition A.2 (Measurability criterion [BK13, pg. 9])**.** *Let $(S, \mathcal{A})$ and $(S', \mathcal{A}')$ be measurable spaces, and let $\mathcal{E}'$ generate $\mathcal{A}'$. Then a mapping $\varphi : S \to S'$ is $\mathcal{A}$-$\mathcal{A}'$-measurable if*

$$\varphi^{-1}(A') \in \mathcal{A} \quad \text{for all} \quad A' \in \mathcal{E}'.$$

Definition A.1 (Weak Measurability)**.** *Let $(\Omega, \mathcal{F})$ be a measurable space and E a Banach space. A function $f : \Omega \to E$ is said to be* weakly measurable *if for every continuous linear functional $\varphi \in E^*$,* *the scalar function*

$$\omega \mapsto \langle \varphi, f(\omega) \rangle$$

is measurable.

Proposition A.3 (Weak Measurability of Measure-Valued Maps)**.** *Let X be a compact Hausdorff space and let $\mathcal{M}(X)$ denote the Banach space of finite Radon measures on X, equipped with the weak* topology induced by the duality with $C_b(X)$. Let*

$$\mu : \Omega \to \mathcal{M}(X)$$

be a map from a measurable space $(\Omega, \mathcal{F})$. Then the following are equivalent:

1. *The map μ is weakly* measurable, i.e., measurable with respect to the weak* topology $\sigma(\mathcal{M}(X), C_b(X))$.*
2. *For every $f \in C_b(X)$, the function*

$$\omega \mapsto \int_X f(x)\, d\mu(\omega)(x)$$

is measurable.

Proof. By the Riesz Representation Theorem, $\mathcal{M}(X) \cong C_b(X)^*$, where each $\mu \in \mathcal{M}(X)$ acts on $f \in C_b(X)$ via

$$\langle \mu, f \rangle = \int_X f\, d\mu.$$

This defines the weak* topology on $\mathcal{M}(X)$, so a map $\mu : \Omega \to \mathcal{M}(X)$ is weakly* measurable if and only if the scalar function

$$\omega \mapsto \langle \mu(\omega), f \rangle = \int_X f\, d\mu(\omega)$$

is measurable for every $f \in C_b(X)$. This is exactly statement 2, so the two are equivalent. $\square$

A.2 Measure

Proposition A.4 (Properties of a Measure [BK13, pg. 20])**.** *For any measure μ and any measurable sets $A, A_1, A_2, \ldots$, there holds:*

(i) ***Monotonicity:*** *$\mu(A_1) \le \mu(A_2)$, if $A_1 \subseteq A_2$,*

(ii) ***σ-subadditivity:***

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \le \sum_{n=1}^{\infty} \mu(A_n),$$

(iii) ***σ-continuity:*** *If $A_n \uparrow A$, then $\mu(A_n) \to \mu(A)$ as $n \to \infty$. If $A_n \downarrow A$ and moreover $\mu(A_1) < \infty$, then $\mu(A_n) \to \mu(A)$ as $n \to \infty$ as well.*

Definition A.2 (Finite Measure)**.** *A measure μ defined on a σ-algebra $\mathcal{F}$ is said to be finite if the measure of the entire space X is finite, that is,*

$$\mu(X) < \infty.$$

Definition A.3 (σ-Finite Measure)**.** *A measure μ defined on a σ-algebra $\mathcal{F}$ is said to be σ-finite if the space X can be decomposed into a countable union of sets $A_n \in \mathcal{F}$, such that each set A_n has finite measure, i.e.,*

$$X = \bigcup_{n=1}^{\infty} A_n \quad \text{and} \quad \mu(A_n) < \infty \text{ for all } n \in \mathbb{N}.$$

Definition A.4. *Let $(S, \mathcal{A})$, $(S', \mathcal{A}')$ be measurable spaces, let $\varphi : S \to S'$ be measurable, and let μ be a measure on $\mathcal{A}$. The measure μ' on S' given by*

$$\mu'(A') = \mu(\varphi^{-1}(A')), \quad A' \in \mathcal{A}',$$

is called the image measure of μ under the mapping φ. We write $\mu' = \varphi(\mu) = \mu \circ \varphi^{-1}$ [BK13, pg. 22].

For the next Proposition we need the notion of $\cap$-stability. We say a generator $\mathcal{E}$ is $\cap$- stable if $E_1, E_2 \in \mathcal{E} \implies E1 \cap E2 \in \mathcal{E}$.

Proposition A.5 (Uniqueness Theorem for Measures, [BK13, Proposition 7.1])**.** *Let $\mathcal{E}$ be a $\cap$-stable generator of a σ-algebra $\mathcal{A}$ on a set S, and let μ and ν be two measures on $\mathcal{A}$. If*

(i) $\mu(E) = \nu(E)$ for every $E \in \mathcal{E}$,

(ii) $\mu(S) = \nu(S) < 1$ or $\mu(E_n) = \nu(E_n) < 1$ for some sets $E_1, E_2, \ldots \in \mathcal{E}$ with $E_n \uparrow S$,

then $\mu = \nu$.

Definition A.5 (Absolute Continuity)**.** *A measure μ on a measurable space $(X, \mathcal{A})$ is said to be* absolutely continuous *with respect to a measure λ, written $\mu \ll \lambda$, if for every measurable set $A \in \mathcal{A}$ with $\lambda(A) = 0$, it follows that $\mu(A) = 0$.*

Remark A.1. *If $\mu \ll \lambda$, and λ is the Lebesgue measure on $\mathbb{R}^n$, then by the* Radon–Nikodym theorem*, there exists a measurable function $g \in L^1(\lambda)$ such that*

$$\mu(A) = \int_A g \, d\lambda \quad \text{for all measurable } A \subseteq X.$$

The function $g = \frac{d\mu}{d\lambda}$ is called the Radon–Nikodym derivative *or the* density *of μ with respect to λ.*

Let $f : X \to \mathbb{R}$ be a measurable function. If $\mu \ll \lambda$ and $\frac{d\mu}{d\lambda} = g$, then

$$\int_X f \, d\mu = \int_X f g \, d\lambda.$$

If g is continuous, then fg is Lebesgue measurable, and:

- If f is integrable with respect to λ, then $fg \in L^1(\lambda)$, so $\int_X fg \, d\lambda$ exists.
- If f is continuous and $g \in L^1(\lambda)$, then $fg \in L^1(\lambda)$, since the product of a continuous and an integrable function is integrable on a compact domain (or under appropriate decay conditions).

A.3 Integral

Proposition A.6 (Monotone Convergence Theorem [BK13, pg. 31])**.** *Let $f_1 \leq f_2 \leq \cdots$ hold for measurable functions $f_1, f_2, \ldots$, and set $f = \sup_{n \geq 1} f_n$. Then*

$$\int f \, d\mu = \lim_{n \to \infty} \int f_n \, d\mu$$

Theorem A.1 (Tonelli's Theorem – Integrability Criterion [Bog07a, Thm. 3.4.5])**.** *Let $f : X \times Y \to [0, \infty]$ be a nonnegative measurable function with respect to the product σ-algebra, where μ and ν are σ-finite measures on X and Y, respectively. Then:*

$$f \in L^1(\mu \otimes \nu) \quad \text{if and only if} \quad \int_Y \left(\int_X f(x, y) \, d\mu(x) \right) d\nu(y) < \infty.$$

Moreover, in that case,

$$\int_{X \times Y} f(x, y) \, d(\mu \otimes \nu)(x, y) = \int_Y \left(\int_X f(x, y) \, d\mu(x) \right) d\nu(y) = \int_X \left(\int_Y f(x, y) \, d\nu(y) \right) d\mu(x).$$

B Proofs and Discussions

B.1 Proof of Lemma 1.1

Lemma B.1. *Let $\{\mu_i\}_{i\in\mathbb{N}}$ be a collection of finite measures, where each μ_i is defined on a corresponding σ-algebra $\mathcal{A}_i$ for each $i \in \mathbb{N}$. Then $\mu = \sum_i \mu_i$ is a measure on the σ-algebra $\mathcal{A} = \bigcap_i \mathcal{A}_i$. The measure space $(S, \mathcal{A}, \mu)$ enjoys the property that for any measurable function $f : S \to \overline{\mathbb{R}}$, we have that*

$$\int f d\mu = \sum_i \int f d\mu_i, \tag{B.13}$$

holds. The symbol $\overline{\mathbb{R}}$ denotes the closure of $\mathbb{R}$. For convenience, and with some abuse of notation, we write $\overline{\mathbb{R}} = [-\infty, \infty]$. Conventions with regard to the symbol ∞ are discussed in Section B.2.

Proof. We shall begin by verifying that μ is indeed a measure on $\mathcal{A}$. Monotonicity and σ-continuity are immediate from the fact that the individual μ_i's are measures (cf. Proposition A.4). The σ-subadditivity property follows from the fact that for any disjoint sequence $\{A_n\}_{n\geq 1}$, where $A_n \in \mathcal{A}$ and $\bigcap_i A_i = \emptyset$, we have that

$$\mu(\cup_n A_n) = \sum_i \sum_n \mu_i(A_n) = \sum_n \sum_i \mu_i(A_n) = \sum_n \mu(A_n),$$

where convergence of the sum follows from the finiteness of the μ_i's and the fact that $A_n \in \mathcal{A}_i$. Next, we show the integration property of the lemma (cf. Eq. (B.13)). For that, let $f \geq 0$ be an elementary function. Then, we have that

$$\int f d\mu = \sum_z z\mu(\{f = z\}) = \sum_z z \sum_i \mu_i(\{f = z\}) = \sum_i \sum_z z\mu_i(\{f = z\}) = \sum_i \int f d\mu_i,$$

where again the convergence of the sum follows from the finiteness of the μ_i's. Now we shall consider for measurable $f \geq 0$, the sequence $\{f_n\}_{n\in\mathbb{N}}$, with

$$f_n = \sum_k \frac{k}{2^n} I_{\{\frac{k}{2^n} \leq f \leq \frac{k+1}{2^n}\}} + 2^n I_{\{f=\infty\}},$$

having the properties that $f_1 \leq f_2 \dots$ and $\sup_{n\geq 1} f_n = f$. This exact same construction was used in [BK13, pg.34]. In fact it can be shown that for any measurable function f there is a sequence of elementary functions, with the desired properties see for example [Bog07a, Corollary 2.1.9]. We then obtain, by Proposition A.6, that

$$\int f\, d\mu = \lim_{n\to\infty} \int f_n\, d\mu = \lim_{n\to\infty} \sum_i \int f_n\, d\mu_i = \sum_i \lim_{n\to\infty} \int f_n\, d\mu_i,$$

holds. The exchange of the sum and limit is also a consequence of Proposition A.6. Here, the sum is considered as an integral with respect to the counting measure (over $\mathbb{N}$), defining the sequence $\{g_n\}_{n\in\mathbb{N}}$, with $g_n(i) = \int f_n\, d\mu_i$, we see that the conditions of Proposition A.6 are satisfied. Explicitly, we have that $g_1 \leq g_2 \leq \dots$ and $\sup_{n\geq 1} g_n(i) = \int f d\mu_i$. Finally, the assertion follows from the fact, that for any integrable function f, we can write $\int f d\mu = \int \max(0, f) d\mu - \int \max(-f, 0) d\mu$. □

B.2 The extended Real Line $\overline{\mathbb{R}}$

The extension $\overline{\mathbb{R}} = [-\infty, \infty]$, i.e the closure of $\mathbb{R}$ is discussed in [BK13, pg. 13]. The following conventions for the symbol ∞ were made:

- $\infty + \infty := \infty$
- $= 0 \cdot \infty := 0$
- $a \cdot \infty := \infty, a > 0$
- $(-1) \cdot \infty := -\infty$

The expressions $\infty - \infty$ and $\frac{\infty}{\infty}$, however remain undefined.

B.3 Discussion on Metric Spaces which are the image of a continuous f under compact interval

For a discussion about compact metric spaces, which are the image of a continuous function over the unit interval $[0, 1]$, the following theorem is really helpful:

Theorem B.1 (Hahn–Mazurkiewicz theorem [JGH12, Theorem 3.30])**.** *A non-empty Hausdorff topological space is a continuous image of the unit interval if and only if it is a compact, connected, locally connected, second-countable space.*

We do not offer a rigorous defintion of second countable, however we mention that a compact metric space, by definition, is second countable. Similar for connect/ locally connected property, we state without proof that the compact metric space $\{[0, 1]^n, d_{[0,1]^n}\}$ possesses these properties for all choices of n and all metrics $d_{[0,1]^n}$.

B.4 Comparison with the ergodicity metric as originally defined in [MM21]

We shall now relate the ergodic framework developed to the metric introduced in [MM21], by [MM21, Eq. (1)] the metric is given by:

$$d^t(B(x, r)) = \frac{1}{Nt} \sum_{j=1}^{N} \int_0^t I_{B(x,r)}(g_j(\tau))\lambda(d\tau),^2 \tag{B.14}$$

where $B(x, r)$ denotes the spherical subsets of U centered at x, formally $B(x, r) = \{y \in U \mid \|y - x\|^2 \leq r\}$. Following Eq. (B.14) we define the system of subsets $\mathcal{E} = \{B(x, r) \mid x \in U, r > 0\}$, where $B(x, r)$ was already defined before. We state that for a compact domain $U \subset \mathbb{R}^n$ any open set can be written as a union of the spherical sets in U and by extension $\mathcal{E}$ generates the Borel σ-algebra of U. At this point we state that all trajectories $g_j : [0, t] \to U$, as they are continuous functions, are $\mathcal{B}$ - $\mathcal{B}^n$ measurable, this is a direct consequence of Proposition A.2, as for continuous functions, the preimages of open sets are again open and hence Borel-measurable. Finally, extending Eq. (B.14) to the Borel sets in U, by considering Borel sets instead of spherical sets, we get the following equation:

$$d^t(B) = \frac{1}{Nt} \sum_{j=1}^{N} \int_0^t I_B(g_j(\tau))\lambda(d\tau), B \in \mathcal{B}^n. \tag{B.15}$$

Using Lemma 3.1 and 1.1, together with the fact that that the probability measure of the uniform distribution over some interval $[0, t]$ is given by $\frac{\lambda(\tau)}{\lambda([0,t])}, \tau \subseteq [0, t]$, measurable, we see that in fact the left hand side of Eq. (1.2) and Eq. (B.15) are in fact equivalent. Finally, as stated before all spherical subset of U are Borel-measurable, therefore Eq. (B.15) is an extension of (B.14).

[2] For better readability we altered the original form, given by: $d^t(x, r) = \frac{1}{Nt} \sum_{j=1}^{N} \int_0^t \chi(x, r)(x_j(\tau))d\tau$

References

[Bax10] Brad Baxter. Conditionally positive functions and p-norm distance matrices, 2010.

[BK13] Martin Brokate and Götz Kersting. *Measure and Integral*. Springer, Berlin, 2013.

[Bog07a] V. I. Bogachev. *Measure Theory, Volume I*. Springer-Verlag, Berlin, 2007.

[Bog07b] V. I. Bogachev. *Measure Theory, Volume II*. Springer-Verlag, Berlin, 2007.

[GS64] I. M. Gel'fand and G. E. Shilov. *Generalized Functions, Volume 1: Properties and Operations*. Academic Press, 1964.

[JGH12] Gail S.Young John G. Hocking. *Topology*. Courier Corporation, 2012.

[Mil76] Stephen C. Milne. Peano curves and smoothness of functions. *Proceedings of the American Mathematical Society*, 58(1):167–173, 1976.

[MM13] Lauren M. Miller and Todd D. Murphey. Trajectory optimization for continuous ergodic exploration on the motion group se(2). In *52nd IEEE Conference on Decision and Control*, pages 4517–4522, 2013.

[MM21] George Mathew and Igor Mezić. Metrics for ergodicity and design of ergodic dynamics for multi-agent systems. 2021.

[Ope25] OpenAI. ChatGPT: Language Model (May 2025 Version). https://chat.openai.com, 2025. Generative AI model developed by OpenAI. Available at: https://chat.openai.com.

[Rud87] Walter Rudin. *Real and Complex Analysis*. McGraw-Hill, New York, 3rd edition, 1987.

[Rud91] Walter Rudin. *Functional Analysis*. McGraw-Hill, New York, 2nd edition, 1991.

[Sag94] Hans Sagan. *Space-Filling Curves*. Springer-Verlag, New York, 1994.

[SGTM25] Max Muchen Sun, Ayush Gaggar, Pete Trautman, and Todd Murphey. Fast ergodic search with kernel functions. *IEEE Transactions on Robotics*, 41:1841–1860, 2025.