



HAL
open science

Entropy minimizing curves with application to flight path design and clustering *

Stéphane Puechmorel, Florence Nicol

► **To cite this version:**

Stéphane Puechmorel, Florence Nicol. Entropy minimizing curves with application to flight path design and clustering *. Entropy, 2016, Special Issue Differential Geometrical Theory of Statistics, 18 (9), pp.337. 10.3390/e18090337. hal-01349675

HAL Id: hal-01349675

<https://hal.science/hal-01349675v1>

Submitted on 28 Jul 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Article

Entropy minimizing curves with application to flight path design and clustering*

Stephane Puechmorel ¹ and Florence Nicol ²

¹ Ecole Nationale de l'Aviation Civile (ENAC)

Dept. SINA/MAIAA

7, Avenue Edouard Belin

F31500 TOULOUSE FRANCE

stephane.puechmorel@enac.fr

² Ecole Nationale de l'Aviation Civile (ENAC)

Dept. SINA/MAIAA

7, Avenue Edouard Belin

F31500 TOULOUSE FRANCE

nicol@recherche.enac.fr

* Correspondence: stephane.puechmorel@enac.fr; Tel.: +33-5-62 17 95 03

Academic Editor: name

Version July 28, 2016 submitted to Entropy; Typeset by L^AT_EX using class file mdpi.cls

1 Air traffic management (ATM) aims at providing companies with a safe and ideally optimal aircraft
 2 trajectory planning. Air traffic controllers act on flight paths in such a way that no pair of aircraft
 3 come closer than the regulatory separation norms. With the increase of traffic, it is expected that
 4 the system will reach its limits in a near future: a paradigm change in ATM is planned with the
 5 introduction of trajectory based operations. In this context, sets of well separated flight paths are
 6 computed in advance, tremendously reducing the number of unsafe situations that must be dealt
 7 with by controllers. Unfortunately, automated tools used to generate such plannings generally
 8 issue trajectories not complying with operational practices or even flight dynamics. In this paper,
 9 a mean of producing realistic air routes from the output of an automated trajectory design tool is
 10 investigated. For that purpose, the entropy of a system of curves is first defined and a mean of
 11 iteratively minimizing it is presented. The resulting curves form a route network that is suitable for
 12 use in a semi-automated ATM system with human in the loop. The tool introduced in this work is
 13 quite versatile and may be applied also to unsupervised classification of curves: an example is given
 14 for the french traffic.

15 1. Introduction

16 Based on recent studies [3], traffic in Europe is expected to grow on an average yearly rate of
 17 2.6%, yielding a net increase of 2 million flights per year at the 2020 horizon. Long term forecast gives
 18 a two fold increase in 2050 over the current traffic, pointing out the need for a paradigm change
 19 in the way flights are managed. Two major framework programs, SESAR (Single European Sky
 20 Air traffic management Research) in Europe and Nextgen in the US have been launched in order
 21 to first investigate potential solutions and to deploy them in a second phase. One of the main
 22 changes that the air traffic management (ATM) system will undergo is a switch from airspace based
 23 to trajectory based operations with a delegation of the separation task to the crews. Within this
 24 framework, trajectories become the basic object of ATM, changing the way air traffic controllers will

*This paper is an extended version of "Entropy minimizing curves with application to automated flight path design published in GSI15

25 be working. In order to alleviate the workload of controllers, trajectories will be planned several
 26 weeks in advance in such a way that close encounters are minimized and ideally removed. For
 27 that purpose, several tools are currently being developed, most of them coming from the field of
 28 robotics [10]. Unfortunately, flight path issued by these algorithms are not tractable for a human
 29 controller and need to be simplified. The purpose of the present work is to introduce an automated
 30 procedure that takes as input a set of trajectories and outputs a simplified one that can be used in
 31 an operational context. Using an entropy associated with a curves system, a gradient descent is
 32 performed in order to reduce it so as to straighten trajectories while avoiding areas with low aircraft
 33 density, thus enforcing route-like behavior. This effect is related to the fact that entropy minimizing
 34 distributions favor concentration.

35 2. Entropy minimizing curves

36 2.1. Motivation

37 As previously mentioned, air traffic management of the future will make an intensive use of
 38 4D trajectories as a basic object. Full automation is a far reaching concept that will probably not
 39 implemented before 2040-2050 and even in such a situation, it will be needed to keep humans in
 40 the loop so as to gain a wide societal acceptance of the concept. Starting from SESAR or Nextgen
 41 initial deployment and aiming towards this ultimate objective, a transition phase with human-system
 42 cooperation will take place. Since ATC controllers are used to a well structured network of routes, it
 43 is advisable to post-process the 4D trajectories issued by automated systems in order to make them as
 44 close as possible to line segments connecting beacons. To perform this task, in an automatic way, flight
 45 paths will be deformed so as to minimize an entropy criterion, that enforces avoidance of low density
 46 area and at the same time penalizes length. Compared to already available bundling algorithms [5]
 47 that tend to move curves to high density areas, this new procedure generates geometrically correct
 48 curves, without excess curvature.

49 Let a set $\gamma_1, \dots, \gamma_N$ of smooth curves be given, that will be aircraft flight paths for the air traffic
 50 application. It will be assumed in the sequel that all curves are smooth mappings from $[0, 1]$ to a
 51 domain Ω of \mathbb{R}^q with everywhere non vanishing derivatives in $]0, 1[$. This last condition allows to
 52 view them as smooth immersions with boundaries and is sound from the application point of view
 53 as aircraft velocities are bounding below by efficiency consideration and ultimately by the stall, and
 54 therefore cannot be vanish expect at the endpoints. In air traffic applications, the dimension of the
 55 state space is generally 2 and sometimes 3 when the evolution of aircraft in the vertical plane is of
 56 interest.

57 The approach taken in this work is first to get a sound definition of spatial density associated
 58 with a curve system, then to derive from it an entropy that will be minimized.

59 2.2. Spatial density of a system of curves

Due to the fact that aircraft positions are acquired through radar measurements, a trajectory is
 known only at discrete sampling times. In the operational context, the sampling period ranges from
 4 to 10 seconds which corresponds roughly to 100-250m traveling distance. Derived from that, a
 classical performance indicator used in ATM is the aircraft density [4], obtained from the sampled
 positions $\gamma_i(t_j)$, $j = 1, \dots, n_i$ on each flight path γ_i , $i = 1, \dots, N$. It is constructed from a partition
 U_k , $k = 1, \dots, P$ of Ω by counting the number of samples occurring in a given U_k then dividing out
 by the total number of samples $n = \sum_{i=1}^N n_i$. More formally, the density d_k in the subset U_k of Ω
 is expressed as:

$$d_k = n^{-1} \sum_{i=1}^N \sum_{j=1}^{n_i} 1_{U_k}(\gamma_i(t_j)) \quad (1)$$

with 1_{U_k} the characteristic function of the set U_k . It seems natural to extend the density obtained from samples to another one based on the trajectories themselves using an integral form:

$$d_k = \lambda^{-1} \sum_{i=1}^N \int_0^1 1_{U_k}(\gamma_i(t)) dt \quad (2)$$

where the normalizing constant λ is chosen so as that d_k is a discrete probability distribution:

$$\lambda = \sum_{k=1}^P \sum_{i=1}^N \int_0^1 1_{U_k}(\gamma_i(t)) dt = \sum_{i=1}^N \int_0^1 \sum_{k=1}^P 1_{U_k}(\gamma_i(t)) dt$$

and since $U_k, k = 1 \dots P$ is a partition:

$$\sum_{k=1}^P 1_{U_k}(\gamma_i(t)) = 1 \quad (3)$$

so that $\lambda = N$.

Density can be viewed as an empirical probability distribution with the U_k considered as bins in an histogram. It is thus natural to extend the above computation so as to give rise to a continuous distribution on Ω . For that purpose, local weighting techniques such as kernel density estimation methods are well-known in nonparametric statistics because they are a useful data driven way to yield continuous density estimation. Many references may be found in the literature as in [11,12]. Given observations, the resulting estimation will be the sum of weights taking into account the distance between the observations and the location x at which the density has to be estimated; the more an observation is close to x , the greater is the weighting. The weights are defined by selecting a summable function centered on the observations, called a kernel, usually denoted by $K: \mathbb{R} \rightarrow \mathbb{R}^+$ in the univariate case, and a smoothed version of the Parzen-Rosenblatt density estimator [8,9] is used. Standard choices for the K function are the ones used for nonparametric kernel estimation like the Epanechnikov function [2]:

$$K: x \mapsto (1 - x^2) 1_{[-1,1]}(x).$$

There exists a large variety of kernel functions and any density function satisfying the normalization condition can be considered so that the estimation is a probability density. Moreover, the kernel function is a symmetric positive function, with a first moment equals to zero and a finite second order moment. In the multivariate case, a multivariate kernel function $\mathcal{K}: \mathbb{R}^q \rightarrow \mathbb{R}^+$ is selected, that can be expressed by means of a real kernel K associated with a norm, denoted by $\|\cdot\|$, in \mathbb{R}^q as follows

$$\mathcal{K}(x) = K(\|x\|), \quad x \in \mathbb{R}^q.$$

The normalization condition becomes

$$\int_{\mathbb{R}^q} \mathcal{K}(x) dx = \int_{\mathbb{R}^q} K(\|x\|) dx = 1.$$

A kernel version of the density is then defined as a mapping d from Ω to $[0, 1]$:

$$d: x \mapsto \frac{\sum_{i=1}^N \int_0^1 K(\|x - \gamma_i(t)\|) dt}{\sum_{i=1}^N \int_{\Omega} \int_0^1 K(\|x - \gamma_i(t)\|) dt dx}. \quad (4)$$

Normalizing the kernel is not mandatory as the normalization occurs with the definition of d . It is nevertheless easier to consider these kind of kernels, as it is done in nonparametric density estimation.

Note that when K is compactly supported, which is the case of the Epanechnikov function and all its relatives, it comes:

$$\int_{\Omega} K(\|x - \gamma_i(t)\|) dx = \int_{\mathbb{R}^q} K(\|x\|) dx$$

provided that Ω contains the set:

$$\{x \in \mathbb{R}^q, \inf_{i=1\dots N, t \in [0,1]} \|x - \gamma_i(t)\| \leq A\}$$

61 where the interval $[-A, A]$ contains the support of K . The case of kernels with unbounded support,
 62 like Gaussian functions, may be dealt with provided $\Omega = \mathbb{R}^q$. In the application considered, only
 63 compactly supported kernels are used, mainly to allow fast machine implementation of the density
 64 computation.

Using the polar coordinates (ρ, θ) and the rotation invariance of the integrand, the relation becomes:

$$\text{Vol}(\mathbb{S}^{q-1}) \int_{\mathbb{R}^+} K(\rho) \rho^{q-1} d\rho = 1$$

which yields a normalizing constant of $2/\pi$ for the Epanechnikov function in dimension 2, instead of the usual $3/4$ in the real case. When the normalization condition is fulfilled, the expression of the density simplifies to:

$$d: x \mapsto N^{-1} \sum_{i=1}^N \int_0^1 K(\|x - \gamma_i(t)\|) dt. \quad (5)$$

65 The normalizing constant is the same as in 2.

66 As an example, one day of traffic over France is considered and pictured on Figure 1 with the
 67 corresponding density map, computed on an evenly spaced grid with a normalized Epanechnikov kernel:

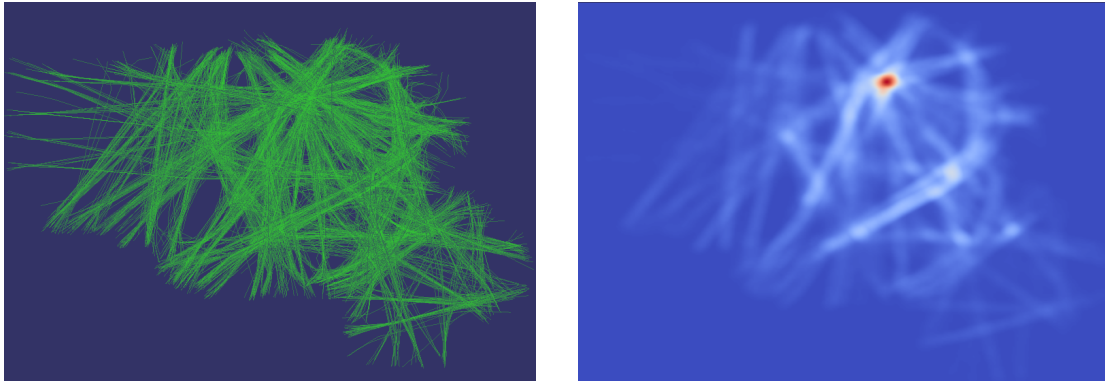


Figure 1. Traffic over France and associated density.

68

Unfortunately, density computed this way suffers a severe flaw for the ATM application: it is not related to the shape of trajectories but also to the time behavior. Formally, it is defined on the set $\mathbf{Imm}([0, 1], \mathbb{R}^q)$ of smooth immersions from $[0, 1]$ to \mathbb{R}^q while the space of primary interest will be the quotient by smooth diffeomorphisms of the interval $[0, 1]$, $\mathbf{Imm}([0, 1], \mathbb{R}^q) / \mathbf{Diff}([0, 1])$. Invariance of the density under the action of $\mathbf{Diff}([0, 1])$ is obtained as in [6] by adding a term related to velocity in the integrals. The new definition of d becomes:

$$\tilde{d}: x \mapsto \frac{\sum_{i=1}^N \int_0^1 K(\|x - \gamma_i(t)\|) \|\gamma_i'(t)\| dt}{\sum_{i=1}^N \int_{\Omega} \int_0^1 K(\|x - \gamma_i(t)\|) \|\gamma_i'(t)\| dt dx}. \quad (6)$$

Assuming again a normalized kernel and letting l_i be the length of the curve γ_i , the expression of the density simplifies to:

$$\tilde{d}: x \mapsto \frac{\sum_{i=1}^N \int_0^1 K(\|x - \gamma_i(t)\|) \|\gamma_i'(t)\| dt}{\sum_{i=1}^N l_i}. \quad (7)$$

69 The new **Diff**-invariant density is pictured on Figure 2 along with the standard density. While the
 70 overall aspect of the plot is similar, one can observe that routes are more apparent the right picture
 71 and that the density peak located above Paris area is of less importance and less symmetric is due to
 72 the fact that near airports, aircraft are slowing down and this effect exaggerates the density with the
 non-invariant definition.

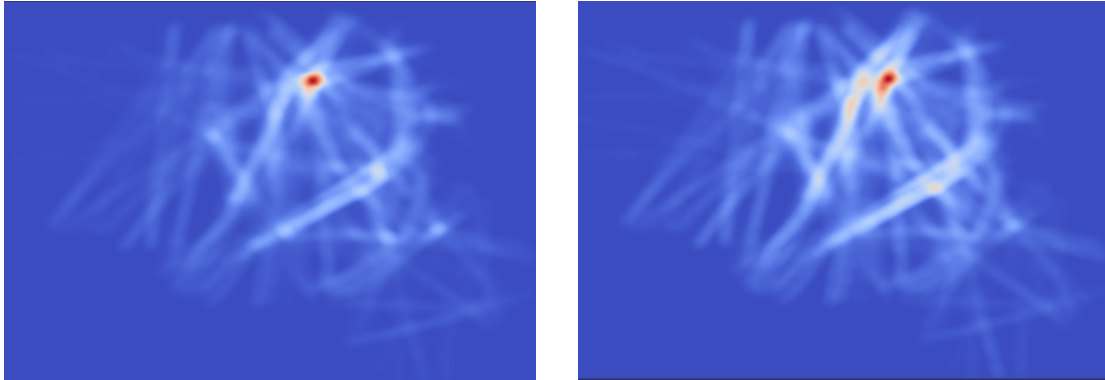


Figure 2. Density (left) and **Diff** invariant density (right) for the 12th February 2013 traffic.

73

74 The extension of the two dimensional defined that way to the general case of curves in an
 75 arbitrary space \mathbb{R}^q is straightforward.

76 2.3. Further properties of the density

77 In this section, curves considered are smooth mapping from the closed interval $[0, 1]$ to \mathbb{R}^q , with
 78 a non-vanishing derivative in $]0, 1[$. All multivariate kernels \mathcal{K} will be assumed smooth, positive,
 79 with a unit integral, and of the form $x \mapsto K(\|x\|)$. However, it is not required that they are compactly
 80 supported unless explicitly stated. All results are presented for the whole space \mathbb{R}^q , but apply almost
 81 verbatim to an open subset.

Definition 1. Let f be a smooth summable mapping from \mathbb{R} to \mathbb{R} . The scaling f_ν of f is defined, for each $\nu > 0$ to be the mapping:

$$f_\nu: x \in \mathbb{R} \mapsto \frac{1}{\nu} f\left(\frac{x}{\nu}\right).$$

82 It is clear that the L^1 -norm of the original mapping is preserved by the scaling. Given a summable
 83 kernel function K from \mathbb{R} to \mathbb{R}^+ , it defines a multivariate kernel \mathcal{K} on \mathbb{R}^q that maps x to $K(\|x\|)$. One
 84 may derive from it a parametrized family of kernels in \mathbb{R} by mapping each ν in $]0, 1[$ to the scaled
 85 kernel K_ν . If the original K is of unit integral, so are all the K_ν .

Proposition 2. Let $\gamma: [0, 1] \rightarrow \mathbb{R}^q$ be a smooth path with non-vanishing derivative in $]0, 1[$. Let $K_\nu, \nu > 0$ be a parametrized family of unit integral kernels. The family of Borel measures μ_ν defined for any Borel set A by:

$$\mu_\nu(A) = \int_A \int_0^1 K_\nu(\|x - \gamma(t)\|) \|\gamma'(t)\| dt dx$$

86 is tight and converges narrowly to the Lebesgue measure on $\gamma([0, 1])$.

Proof. Let $\epsilon > 0$ be given. By the summability of K , it exists a positive real number r such that:

$$\int_{\mathbb{R}^q - B(0,r)} K(\|x\|) dx < \epsilon$$

with $B(0,r)$ the open ball of radius r centered at the origin. Since $B(0,r) \subset B(0,rv^{-1})$ for $v > 0$, the same holds for all the family K_v . Let $B(0,M)$ be an open ball containing $\gamma([0,1])$. Then:

$$\mu_v \left(\mathbb{R}^q - \overline{B(0,M+r)} \right) = \int_{\mathbb{R}^q - B(0,M+r)} \int_0^1 K_v(\|x - \gamma(t)\|) \|\gamma'(t)\| dt dx \quad (8)$$

$$= \int_0^1 \int_{\mathbb{R}^q - B(0,M+r)} K_v(\|x - \gamma(t)\|) \|\gamma'(t)\| dx dt \quad (9)$$

$$\leq \epsilon \int_0^1 \|\gamma'(t)\| dt = \epsilon l(\gamma) \quad (10)$$

87 where $l(\gamma)$ denotes the length of γ . This proves the tightness of the family K_v .

Let $f: \mathbb{R}^q \rightarrow \mathbb{R}$ be a bounded continuous mapping. It comes:

$$I_v(f) = \int_{\mathbb{R}^q} \int_0^1 K_v(\|x - \gamma(t)\|) f(x) \|\gamma'(t)\| dt dx = \int_0^1 \int_{\mathbb{R}^q} K_v(\|x - \gamma(t)\|) f(x) \|\gamma'(t)\| dx dt \quad (11)$$

$$= \int_0^1 \int_{\mathbb{R}^q} K(\|x\|) f(xv + \gamma(t)) \|\gamma'(t)\| dx dt \quad (12)$$

and since f is bounded, the dominated convergence theorem shows that:

$$\lim_{v \rightarrow 0} I_v(f) = \int_0^1 f(\gamma(t)) \|\gamma'(t)\| dt$$

88 proving the second part of the claim. \square

The density in 7 is for a single curve of the form $d(x) = l(\gamma)^{-1} \int_0^1 K(\|x - \gamma(t)\|) \|\gamma'(t)\| dt$ with $l(\gamma)$ the length of the curve γ . It is invariant under change of parameter, and can be written in a more concise way as:

$$\int_0^1 K(\|x - \gamma(\eta)\|) \|d\eta \quad (13)$$

89 where η is the arclength times $l(\gamma)^{-1}$.

90 This form allows a simple probabilistic interpretation of the density d : if a point u is drawn on
91 the curve γ according to a uniform distribution and independently a vector v in \mathbb{R}^q with a density \mathcal{K}
92 (the multivariate kernel corresponding to K), then the density of $x = u + v$ is given by the equation
93 13.

Proposition 3. *If the multivariate kernel \mathcal{K} has a finite second moment, that is the univariate kernel K is such that:*

$$M = \int_{\mathbb{R}^+} r^{q+1} K(r) dr < +\infty$$

then the Wasserstein distance between the densities d_1, d_2 associated to smooth curves γ_1, γ_2 is bounded by:

$$2\text{Vol}(\mathbb{S}^{q-1})M + D(\gamma_1, \gamma_2)$$

with :

$$D(\gamma_1, \gamma_2) = \int_0^1 \|\gamma_1(\eta) - \gamma_2(\eta)\|^2 d\eta$$

94 where each curve is parametrized by the scaled arclength as in 13.

Proof. Let us consider the plan [1] given by the density:

$$d: (x, y) \mapsto \int_0^1 K(\|x - \gamma_1(\eta)\|)K(\|y - \gamma_2(\eta)\|)d\eta$$

where each curve is parametrized by the scaled arclength. The associated transport cost is given by:

$$C = \int_{\mathbb{R}^q \times \mathbb{R}^q} \|x - y\|^2 \int_0^1 K(\|x - \gamma_1(\eta)\|) K(\|y - \gamma_2(\eta)\|) d\eta dx dy$$

letting $u = y - x$ and using Fubini gives:

$$C = \int_0^1 \int_{\mathbb{R}^q} K(\|x - \gamma_1(\eta)\|) \int_{\mathbb{R}^q} \|u\|^2 K(\|u + x - \gamma_2(\eta)\|) du dx d\eta.$$

The inner term can be written as:

$$\begin{aligned} \int_{\mathbb{R}^q} \|u\|^2 K(\|u + x - \gamma_2(\eta)\|) du &= \int_{\mathbb{R}^q} \|u + \gamma_2(\eta) - x\|^2 K(\|u\|) du & (14) \\ &= \int_{\mathbb{R}^q} \|u\|^2 K(\|u\|) du + 2\langle \gamma_2(\eta) - x, \int_{\mathbb{R}^q} u K(\|u\|) du \rangle + \|\gamma_2(\eta) - x\|^2. & (15) \end{aligned}$$

The integral:

$$\int_{\mathbb{R}^q} u K(\|u\|) du$$

is 0 and, using spherical coordinates:

$$\int_{\mathbb{R}^q} \|u\|^2 K(\|u\|) du = \int_{\mathbb{R}^+} r^{q+1} K(r) \int_{\mathbb{S}^{q-1}} d\sigma dr = \text{Vol}(\mathbb{S}^{q-1})M$$

with $M = \int_{\mathbb{R}^+} r^{q+1} K(r)$. Putting back this value in the expression of the cost gives:

$$C = \text{Vol}(\mathbb{S}^{q-1})M \int_0^1 \int_{\mathbb{R}^q} K(\|x - \gamma_1(\eta)\|) dx d\eta + \int_0^1 \int_{\mathbb{R}^q} K(\|x - \gamma_1(\eta)\|) \|\gamma_2(\eta) - x\|^2 dx d\eta \quad (16)$$

$$= \text{Vol}(\mathbb{S}^{q-1})M + \int_0^1 \int_{\mathbb{R}^q} K(\|x - \gamma_1(\eta)\|) \|\gamma_2(\eta) - x\|^2 dx d\eta \quad (17)$$

$$= \text{Vol}(\mathbb{S}^{q-1})M + \int_0^1 \int_{\mathbb{R}^q} K(\|x\|) \|\gamma_2(\eta) - \gamma_1(\eta) + x\|^2 dx d\eta. \quad (18)$$

Finally:

$$\int_{\mathbb{R}^q} K(\|x\|) \|\gamma_2(\eta) - \gamma_1(\eta) + x\|^2 dx = \int_{\mathbb{R}^q} K(\|x\|) \|\gamma_2(\eta) - \gamma_1(\eta)\|^2 dx d\eta \quad (19)$$

$$+ 2\langle \gamma_2(\eta) - \gamma_1(\eta), \int_{\mathbb{R}^q} x K(\|x\|) dx \rangle \quad (20)$$

$$+ \text{Vol}(\mathbb{S}^{q-1})M. \quad (21)$$

As before the middle term vanishes and the first one integrates to:

$$\int_0^1 \|\gamma_2(\eta) - \gamma_1(\eta)\|^2 d\eta$$

so that:

$$C = 2\text{Vol}(\mathbb{S}^{q-1})M + \int_0^1 \|\gamma_2(\eta) - \gamma_1(\eta)\|^2 d\eta.$$

96 This result indicates that the densities associated to curves γ_1, γ_2 using the smoothing process
 97 described above cannot be too far (with respect to the Wasserstein distance) one from each other if the
 98 geometric L^2 distance $D(\gamma_1, \gamma_2)$ is small. In fact, the upper bound in Proposition 3 can be interpreted
 99 as the cost of moving the smoothed density around γ_1 to the uniform distribution on the curve, then
 100 move γ_1 to γ_2 keeping points with equal scaled arclength in correspondence, and finally move the
 101 uniform distribution on γ_2 to the smoothed density.

Having a density at hand, the entropy of the system of curves $\gamma_1, \dots, \gamma_N$ is defined the usual way as:

$$E(\gamma_1, \dots, \gamma_N) = - \int_{\Omega} \tilde{d}(x) \log(\tilde{d}(x)) dx.$$

102 The entropy is dependent on the particular choice of the kernel K . As mentioned before, it is a
 103 common practice in the field of non-parametric statistics to introduce a tuning parameter $\nu > 0$ in
 104 the kernel, called bandwidth, so that is it expressed as a scaled version $K = f_{\nu}$ of a given function
 105 $f: \mathbb{R}^+ \rightarrow \mathbb{R}^+$. The value of ν is the most influential parameter in the estimation of the density and
 106 must be selected carefully. For curves clustering applications, it is defined by the desired interaction
 107 length: if ν tends to 0, the curves will behave as independent objects while on the other end of the scale
 108 very high bandwidth will tend to remove the influence of the curves themselves. For the moment, no
 109 automated mean of finding an optimal ν was used, although it will be part of a future work.

110 2.4. Minimizing the entropy

In order to fulfill the initial requirement of finding bundles of curve segments as straight as possible, one seeks after the system of curves minimizing the entropy $E(\gamma_1, \dots, \gamma_N)$, or equivalently maximizing:

$$\int_{\Omega} \tilde{d}(x) \log(\tilde{d}(x)) dx.$$

111 The reason why this criterion gives the expected behavior will become more apparent after derivation
 112 of its gradient at the end of this part. Nevertheless, when considering a single trajectory it is intuitive
 113 that the most concentrated density distribution is obtained with a straight segment connecting the
 114 endpoints: this point will be made rigorous later.

Letting ϵ be a perturbation of the curve γ_j such that $\epsilon(0) = \epsilon(1) = 0$, the first order expansion of $-E(\gamma_1, \dots, \gamma_N)$ will be computed in order to get a maximizing displacement field, analogous to a gradient ascent¹ in the finite dimensional setting. The notation:

$$\frac{\partial F}{\partial \gamma_j}$$

will be used in the sequel to denote the derivative of a function F of the curve γ_j in the sense that for a perturbation ϵ :

$$F(\gamma_j + \epsilon) = F(\gamma_j) + \frac{\partial F}{\partial \gamma_j}(\epsilon) + o(\|\epsilon\|_2).$$

First of all, please note that since \tilde{d} has integral 1 over the domain Ω :

$$\int_{\Omega} \frac{\partial \tilde{d}(x)}{\partial \gamma_j}(\epsilon) dx = 0$$

so that:

$$- \frac{\partial}{\partial \gamma_j} E(\gamma_1, \dots, \gamma_N)(\epsilon) = \int_{\Omega} \frac{\partial \tilde{d}(x)}{\partial \gamma_j}(\epsilon) \log(\tilde{d}(x)) dx. \quad (22)$$

¹ Choice has been made to maximize the opposite of the entropy, so that the algorithm will be a gradient ascent one.

Starting from the expression of \tilde{d} given in equation 7, the first order expansion of \tilde{d} with respect to the perturbation ϵ of γ_j is obtained as a sum of a term coming from the numerator:

$$\int_0^1 K(\|x - \gamma_j(t)\|) \|\gamma_j'(t)\| dt. \quad (23)$$

and a second one coming from the length of γ_j in the denominator. This last term is obtained from the usual first order variation formula of a curve length:

$$\begin{aligned} & \int_{[0,1]} \|\gamma_j'(t) + \epsilon'(t)\| dt = \\ & \int_{[0,1]} \|\gamma_j'(t)\| dt + \int_{[0,1]} \left\langle \frac{\gamma_j'(t)}{\|\gamma_j'(t)\|}, \epsilon'(t) \right\rangle dt + o(\|\epsilon\|_2). \end{aligned}$$

Using an integration by part, the first order term can be written as:

$$\int_{[0,1]} \left\langle \frac{\gamma_j'(t)}{\|\gamma_j'(t)\|}, \epsilon'(t) \right\rangle dt = \quad (24)$$

$$- \int_{[0,1]} \left\langle \left(\frac{\gamma_j''(t)}{\|\gamma_j'(t)\|} \right)_{\mathcal{N}}, \epsilon(t) \right\rangle dt \quad (25)$$

with:

$$\left(\frac{\gamma_j''(t)}{\|\gamma_j'(t)\|} \right)_{\mathcal{N}} = \frac{\gamma_j''(t)}{\|\gamma_j'(t)\|} - \frac{\gamma_j'(t)}{\|\gamma_j'(t)\|} \left\langle \frac{\gamma_j'(t)}{\|\gamma_j'(t)\|}, \frac{\gamma_j''(t)}{\|\gamma_j'(t)\|} \right\rangle$$

the normal component of:

$$\frac{\gamma_j''(t)}{\|\gamma_j'(t)\|}.$$

- 115 Please note that when dealing with planar curves (i.e. with values in \mathbb{R}^2), it is $\kappa_j(t)N_j(t)$ with κ_j (resp
116 N_j) the curvature (resp. the unit normal vector) of γ_j .

The integral in 23 can be expanded in a similar fashion. Using as above the notation $(\cdot)_{\mathcal{N}}$ for normal components, the first order term is obtained as:

$$\int_{[0,1]} \left\langle \left(\frac{\gamma_j(t) - x}{\|\gamma_j(t) - x\|} \right)_{\mathcal{N}}, \epsilon(t) \right\rangle K'(\|\gamma_j(t) - x\|) \|\gamma_j'(t)\| dt \quad (26)$$

$$- \int_{[0,1]} \left\langle \left(\frac{\gamma_j''(t)}{\|\gamma_j'(t)\|} \right)_{\mathcal{N}}, \epsilon(t) \right\rangle K(\|\gamma_j(t) - x\|) dt. \quad (27)$$

From the expressions in 26 and 24, the first order variation of the entropy is:

$$\frac{1}{\sum_{i=1}^N l_i} \left(\int_{[0,1]} \left\langle \int_{\Omega} \left(\frac{\gamma_j(t) - x}{\|\gamma_j(t) - x\|} \right)_{\mathcal{N}} K'(\|\gamma_j(t) - x\|) \log(\tilde{d}(x)) dx, \epsilon(t) \right\rangle \|\gamma_j'(t)\| dt \right) \quad (28)$$

$$- \int_{[0,1]} \left(\int_{\Omega} K(\|\gamma_j(t) - x\|) \log(\tilde{d}(x)) dx \right) \left\langle \left(\frac{\gamma_j''(t)}{\|\gamma_j'(t)\|} \right)_{\mathcal{N}}, \epsilon(t) \right\rangle dt \quad (29)$$

$$+ \left(\int_{\Omega} \tilde{d}(x) \log(\tilde{d}(x)) dx \right) \int_{[0,1]} \left\langle \left(\frac{\gamma_j''(t)}{\|\gamma_j'(t)\|} \right)_{\mathcal{N}}, \epsilon(t) \right\rangle dt. \quad (30)$$

As expected, only moves normal to the trajectory will change at first order the value of the criterion: the displacement of the curve γ_j will thus be performed at t in the normal bundle to γ_j and is given, up to the $(\sum_{i=1}^N l_i)^{-1}$ term by:

$$\int_{\Omega} \left(\frac{\gamma_j(t) - x}{\|\gamma_j(t) - x\|} \right)_{\mathcal{N}} K'(\|\gamma_j(t) - x\|) \log(\tilde{d}(x)) dx \|\gamma_j'(t)\| \quad (31)$$

$$- \left(\int_{\Omega} K(\|\gamma_j(t) - x\|) \log(\tilde{d}(x)) dx \right) \left(\frac{\gamma_j''(t)}{\|\gamma_j'(t)\|} \right)_{\mathcal{N}} \quad (32)$$

$$+ \left(\int_{\Omega} \tilde{d}(x) \log(\tilde{d}(x)) dx \right) \left(\frac{\gamma_j''(t)}{\|\gamma_j'(t)\|} \right)_{\mathcal{N}}. \quad (33)$$

117 The first term in the expression will favor moves towards areas of high density, while the second
 118 and third one are moving along normal vector and will straighten trajectory. This last point can be
 119 enlightened by considering the case of a single planar curve with fixed endpoints.

120 **Proposition 4.** Let a, b be fixed points in \mathbb{R}^2 and K be a kernel as in 7. The segment $[a, b]$ is a critical point for
 121 the entropy associated with the curve system in \mathbb{R}^2 consisting of single smooth paths with endpoints a, b .

Proof. Let the segment $[a, b]$ be parametrized as $\gamma: t \in [0, 1] \mapsto a + tv$ with v the vector $(b - a)$. Starting with the expression 31, it is clear that the second and third terms occurring in the formula will vanish as the second derivative of γ is 0. Let u be the unit normal vector to γ . Any point x in \mathbb{R}^2 can be written as $x = a + \theta v + \zeta u$, $\theta, \zeta \in \mathbb{R}$, so that $\gamma(t) - x = (t - \theta)v - \zeta u$ and $\|\gamma(t) - x\| = \sqrt{(t - \theta)^2 \|b - a\|^2 + \zeta^2}$. The change of variables $x \rightarrow (\theta, \zeta)$ has Jacobian $\|v\| = \|b - a\|$. For a fixed $t \in [0, 1]$, it comes:

$$\int_{\mathbb{R}^2} \left(\frac{\gamma(t) - x}{\|\gamma(t) - x\|} \right)_{\mathcal{N}} K'(\|\gamma(t) - x\|) \log(\tilde{d}(x)) dx \|\gamma'(t)\| = \quad (34)$$

$$\|b - a\|^2 \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{-\zeta}{\sqrt{(t - \theta)^2 \|b - a\|^2 + \zeta^2}} K' \left(\sqrt{(t - \theta)^2 \|b - a\|^2 + \zeta^2} \right) \log(\tilde{d}(\theta, \zeta)) d\zeta d\theta. \quad (35)$$

The density \tilde{d} for the γ curve is expressed in ζ, θ coordinates as:

$$\int_{[0,1]} K \left(\sqrt{(t - \theta)^2 \|b - a\|^2 + \zeta^2} \right) dt$$

and is an even function in ζ . The same is true for $K'(\|\gamma(t) - x\|)$. Finally, the mapping:

$$\zeta \mapsto \frac{-\zeta}{\sqrt{(t - \theta)^2 \|b - a\|^2 + \zeta^2}}$$

122 is odd for a fixed θ so that the whole integrand is odd as a function of ζ . By the Fubini theorem,
 123 integrating first in ζ will therefore yield a vanishing integral, proving the assertion. \square

124 The result still holds in \mathbb{R}^q , the only different aspect being that x is now expanded as $x = a +$
 125 $\theta v + \sum_{i=1}^{q-1} \zeta_i u_i$ with $u_i, i = 1, \dots, q - 1$ an orthonormal basis of the orthogonal complement of $\mathbb{R}v$ in
 126 \mathbb{R}^q . Rewriting $\gamma(t) - x = (t - \theta)v - \sum_{i=1}^{q-1} \zeta_i u_i$ and $\|\gamma(t) - x\| = \sqrt{(t - \theta)^2 \|b - a\|^2 + \sum_{i=1}^{q-1} \zeta_i^2}$, the
 127 same parity argument can be applied on any of the components $\zeta_i, i = 1, \dots, q - 1$, showing that the
 128 integral is vanishing.

129 The effect of curve straightening is present when minimizing the entropy of a whole curve
 130 system, but is counterbalanced by the gathering effect. Depending on the choice of the kernel
 131 bandwidth, one or the other effect is dominant: straightening is preminent for low values, being

132 the only remaining effect in the limit, while gathering dominates at high bandwidths. For the air
 133 traffic application, a rule of the thumb is to take 2-3 times the separation norm as an effective support
 134 for the kernel. Using an adaptive bandwidth may be of some interest also: starting with medium to
 135 high values favors curves gathering, then gradually reducing it will straighten the trajectories.

Using the scaled arclength in the entropy gives an equivalent but somewhat easier to interpret result. Starting with the expression 7, that takes in this case the form:

$$\tilde{d}: x \mapsto \frac{\sum_{i=1}^N l_i \int_0^1 K(\|x - \gamma_i(\eta)\|) d\eta}{\sum_{i=1}^N l_i}. \quad (36)$$

136 Let $i \in \{1, \dots, N\}$ be fixed. An admissible variation of the curve γ_i is a smooth mapping from
 137 $] - a, a[\times [0, 1]$ to \mathbb{R}^q , with $a > 0$ satisfying the following properties:

- 138 (a) $\forall \eta \in [0, 1], \phi(0, \eta) = \gamma_i(\eta)$.
 139 (b) $\forall (t, \eta) \in] - a, a[\times] 0, 1[, \|\partial_\eta \phi(t, \eta)\| = l_\phi(t)$ with $l_\phi(t)$ the length of the curve $\eta \mapsto \phi(t, \eta)$.
 140 (c) $\forall t \in] - a, a[, \phi(t, 0) = \gamma_i(0), \phi(t, 1) = \gamma_i(1)$.

Taking the derivative with respect to t at 0 of equation (b) yields:

$$\langle \partial_t \partial_\eta \phi(0, \eta), \partial_\eta \phi(0, \eta) \rangle = \partial_t l_\phi(0) l_i.$$

Letting $T(\eta)$ be the unit tangent vector to γ_i at η , and noting that $\partial_\eta \phi(0, \eta) = l_i T(\eta)$, it comes:

$$\langle \partial_t \partial_\eta \phi(0, \eta), T(\eta) \rangle = \partial_t l_\phi(t). \quad (37)$$

141 This relation puts a constraint on the variation of the tangential component of the curve derivative
 142 and shows that it has to be constant in η .

Proposition 5. Let D be the mapping from $] - a, a[\times \mathbb{R}^q$ to \mathbb{R}^+ defined by:

$$D: (t, x) \mapsto \frac{\sum_{j=1, j \neq i}^N l_j \int_0^1 K(\|x - \gamma_j(\eta)\|) d\eta + \int_0^1 K(\|x - \phi(t, \eta)\|) d\eta}{\sum_{j=1}^N l_j}.$$

where η refers collectively to the scaled arclength parameter for each curve. The partial derivative $\partial_t D(0, x)$ is given by :

$$\partial_t D(0, x) = \frac{l_i}{\sum_{j=1}^N l_j} \int_0^1 \left\langle \frac{\gamma_i(\eta) - x}{\|\gamma_i(\eta) - x\|}, \partial_t \phi(0, \eta) \right\rangle K'(\|\gamma_i(\eta) - x\|) d\eta.$$

The proof is straightforward and is omitted. From Proposition 5 the variation of the entropy is derived:

$$\partial_t E = - \int_{\mathbb{R}^q} \frac{l_i}{\sum_{j=1}^N l_j} \int_0^1 \left\langle \frac{\gamma_i(\eta) - x}{\|\gamma_i(\eta) - x\|}, \partial_t \phi(0, \eta) \right\rangle K'(\|\gamma_i(\eta) - x\|) d\eta dx. \quad (38)$$

This relation is equivalent to 28: it can be seen by splitting the terms into a normal and a tangential component. The first one yields:

$$- \int_{\mathbb{R}^q} \frac{l_i}{\sum_{j=1}^N l_j} \int_0^1 \left\langle \left(\frac{\gamma_i(\eta) - x}{\|\gamma_i(\eta) - x\|} \right)_{\mathcal{N}}, (\partial_t \phi(0, \eta))_{\mathcal{N}} \right\rangle K'(\|\gamma_i(\eta) - x\|) d\eta dx.$$

For the tangential part, the starting point is the relation:

$$\partial_\eta (K(\|\phi(0, \eta) - x\|) T(\eta)) = l_i K'(\|\phi(0, \eta) - x\|) \left\langle \frac{\phi(0, \eta) - x}{\|\phi(0, \eta) - x\|}, T(\eta) \right\rangle T(\eta) \quad (39)$$

$$+ K(\|\phi(0, \eta) - x\|) \partial_\eta T(\eta). \quad (40)$$

where the subscript \mathcal{T} stands for tangential component. It comes:

$$\frac{l_i}{\sum_{j=1}^N l_j} \int_0^1 \left\langle \left(\frac{\gamma_i(\eta) - x}{\|\gamma_i(\eta) - x\|} \right)_{\mathcal{T}}, (\partial_t \phi(0, \eta))_{\mathcal{T}} \right\rangle K'(\|\gamma_i(\eta) - x\|) d\eta dx = \quad (41)$$

$$\frac{l_i}{\sum_{j=1}^N l_j} \int_0^1 \langle \partial_\eta (K(\|\phi(0, \eta) - x\|) T(\eta)), (\phi(0, \eta))_{\mathcal{T}} \rangle d\eta dx \quad (42)$$

$$- \frac{l_i}{\sum_{j=1}^N l_j} \int_0^1 \langle K(\|\phi(0, \eta) - x\|) \partial_\eta T(\eta), (\phi(0, \eta))_{\mathcal{T}} \rangle d\eta dx. \quad (43)$$

With an integration by parts, the first integral in the right hand side becomes:

$$- \frac{l_i}{\sum_{j=1}^N l_j} \int_0^1 \langle K(\|\phi(0, \eta) - x\|) T(\eta), \partial_\eta (\partial_t \phi(0, \eta))_{\mathcal{T}} \rangle d\eta dx = \quad (44)$$

$$- \frac{l_i}{\sum_{j=1}^N l_j} \int_0^1 K(\|\phi(0, \eta) - x\|) \partial_t l_\phi(0) d\eta dx. \quad (45)$$

Gathering terms, the expression 28 is recovered. As expected, only the normal components enter the relation, but it has to be noted that the tangential component of $\partial_t \phi(0, \eta)$ is not arbitrary and can be deduced from 37. The gradient with respect to the i -th curve is obtained from the expression of the entropy variation and can be written in its simplest form as:

$$\frac{l_i}{\sum_{j=1}^N l_j} \int_{\mathbb{R}^q} \int_0^1 \frac{\gamma_i(\eta) - x}{\|\gamma_i(\eta) - x\|} K'(\|\gamma_i(\eta) - x\|) d\eta \log \tilde{d}(x) dx. \quad (46)$$

143 where \tilde{d} is the estimated spatial density. One must keep in mind the constraint on $\partial_t \phi(0, \eta)$ that is
144 hidden within the apparent simplicity of the expression.

145 3. Numerical implementation

The two formulations 31, 46 of the gradient may be used. The first one is more complicated but does not require any additional constraint to be taken into account. The second one cannot be applied readily as the tangential component must comply with relation 37. In both cases, **F**: it is needed to evaluate a spatial integral, which may yield to prohibitive computational time, especially in high dimensions. In the air traffic application, only planar or 3D curves are considered, greatly simplifying the problem. Nevertheless, the performance of the algorithms is still a concern and the choice made was to replace the spatial integral by a discrete sum over an evenly spaced grid. From now, it is assumed that all curves are planar, so that the ambient space for the spatial density \tilde{d} is \mathbb{R}^2 . Going back to the expression of \tilde{d} given by 7, a first step is to replace the integral over t by a discrete sum. In practice, curves are described by a sequence of sampled points $\gamma_i(t_{ij})$ where the sampling times t_{ij} will be assumed to be identical for all curves. This assumption is not satisfied in the air traffic application so that a pre-processing step must be taken before the actual entropy minimization stage. It will not be described here as any standard interpolation procedure can be applied with negligible differences on the final result. To obtain the results presented here, a cubic spline smoother was used. Since the sampling times are assumed to be the same for all trajectories, the double subscript will be dropped, so that the samples on each trajectory will be denoted as $\gamma_{ij} = \gamma_i(t_j)$. It is further assumed that the derivative $\gamma'_{ij} = \gamma'_i(t_j)$ is available, most of the time through a numerical approximation. Given a quadrature formula on $[0, 1]$ with points $t_j, j = 1, \dots, m$ and weights $w_j, j = 1, \dots, m$, the density may be approximated at $x \in \mathbb{R}^2$ by:

$$\tilde{d}(x) = \frac{1}{\sum_{i=1}^N l_i} \sum_{i=1}^N \sum_{j=1}^m w_j K(\|x - \gamma_{ij}\|) \|\gamma'_{ij}\|. \quad (47)$$

where the lengths $l_i, i = 1, \dots, N$ are also obtained with the same quadrature rule:

$$l_i = \sum_{j=1}^m w_j \|\gamma'_{ij}\|.$$

When γ'_{ij} is computed in a numerical way, it may be expressed as:

$$\gamma'_{ij} = \sum_{k=1}^m \tilde{w}_{jk} \gamma_{i,k}.$$

where the weights \tilde{w}_{jk} are often obtained through the application of the Lagrange interpolation formula to ensure exactness on polynomials up to a given degree. In a more compact form, it can be written in matrix form as:

$$\begin{pmatrix} \gamma'_{i1} \\ \vdots \\ \gamma'_{iq} \end{pmatrix} = \tilde{W} \begin{pmatrix} \gamma_{i1} \\ \vdots \\ \gamma_{iq} \end{pmatrix}$$

where the matrix \tilde{W} has entries the weights \tilde{w}_{jk} . The cost of evaluating \tilde{d} at a single point is in $o(Nm)$, with the kernel evaluation being dominant. When dealing with points in \mathbb{R}^2 or \mathbb{R}^3 and compactly supported kernels a simple trick greatly reduces the time needed to compute \tilde{d} . First of all, the domain of interest is discretized on a evenly spaced grid, so that points of evaluation of the density \tilde{d} are its vertices $x_{ij}, i = 1, \dots, n_x, j = 1, \dots, n_y$. The grid step δ_x (resp. δ_y) in the first (resp. second) coordinate is the difference between any two adjacent vertices $\delta_x = x_{i+1,j} - x_{i,j}$ (resp. $\delta_y = x_{i,j+1} - x_{i,j}$ (most of the time, $\delta_x = \delta_y$). Since the expression 47 is linear, the computation can be performed by accumulating values $K(\|x_{kl} - \gamma_{ij}\|) \|\gamma'_{ij}\|$ for a fixed couple (i, j) , where only the points x_{kl} close enough to γ_{ij} are considered. In fact, the evaluation can be written as a 2D discrete convolution:

$$\tilde{d}(x_{kl}) = \sum_{i=1\dots N, j=1\dots m} w_j K(\|x_{kl} - \gamma_{ij}\|) \|\gamma'_{ij}\|. \quad (48)$$

When the support of K is small compared to the overall spatial domain, a lot of computation is saved using this procedure. Furthermore, it can be thought as 2D filtering, so that highly efficient algorithms coming from the field of image processing can be applied: in particular, computing the density on a graphics processing unit (GPU) is straightforward and allows to decrease the computational time by at least a factor ten. When dealing with the scaled arclength, the derivative term is not present, an a factor of l_i appears in from of the integral. The discrete version becomes:

$$\tilde{d}(x_{kl}) = \sum_{i=1\dots N, j=1\dots m} l_i w_j K(\|x_{kl} - \gamma_{ij}\|) \quad (49)$$

146 where $\gamma_{ij} = \gamma_i(\eta_j)$, η_j being in correspondence with t_j . Please note that the quadrature weights must
 147 be adapted to the abscissa $\eta_j, j = 1 \dots m$ and not to the $t_j, j = 1 \dots m$. Therefore, it is advisable
 148 to resample the curves so that the points $\eta_j, j = 1 \dots m$ are for example evenly spaced or of the
 149 Gauss-Lobatto form. The former was chosen for the experiments due to its ease of implementation,
 150 although the second form is probably more efficient from a numerical point of view and will be
 151 investigated in a second stage.

Having the density at hand, the gradient of the entropy with respect to the points $\gamma_{ij}, i = 1 \dots N, j = 1 \dots m$ can be easily computed using a straightforward application of the formula 31. When dealing with planar curves, a simplification occurs for the second derivative term since for a smooth curve γ_j :

$$\left(\frac{\gamma_j''(t)}{\|\gamma_j'(t)\|} \right)_{\mathcal{N}} = \kappa(t) N(t).$$

152 where κ is the curvature and N the unit normal vector. These quantities may be computed using
 153 numerical differentiation, but a coarse approximation based on the rotation rate of the vectors $\gamma_{i,j+1} -$
 154 $\gamma_{i,j}$, $\gamma_{i,j+2} - \gamma_{i,j+1}$ works well in many cases.

155 The case of scaled arclength parametrization needs some extra attention, due to the condition
 156 on the tangential component. The simplest approach is to move the points γ_{ij} according to an
 157 unconstrained gradient, then to re-sample the obtained curve so as get adjusted γ_{ij} that correspond
 158 to the abscissa $\eta_j, j = 1 \dots m$.

159 In a numerical implementation, the scaling factor in front of the whole expression may be
 160 dropped due to the fact that all gradient-based algorithms will use an automatically tuned step
 161 length. As usual with gradient algorithms, one must carefully select the step taking in the maximizing
 162 direction in order to avoid divergence. A simple fixed step strategy was first applied and gives
 163 satisfactory results on small datasets. A safer approach is to adapt the step size so as to ensure
 164 a sufficient decrease of the entropy. Due to the potentially huge dimension of the search space,
 165 this procedure has to be simple enough. An approximate quadratic search [13] was used in final
 166 implementation.

The procedure applied to one day of traffic over France yields the picture of Figure 3. As

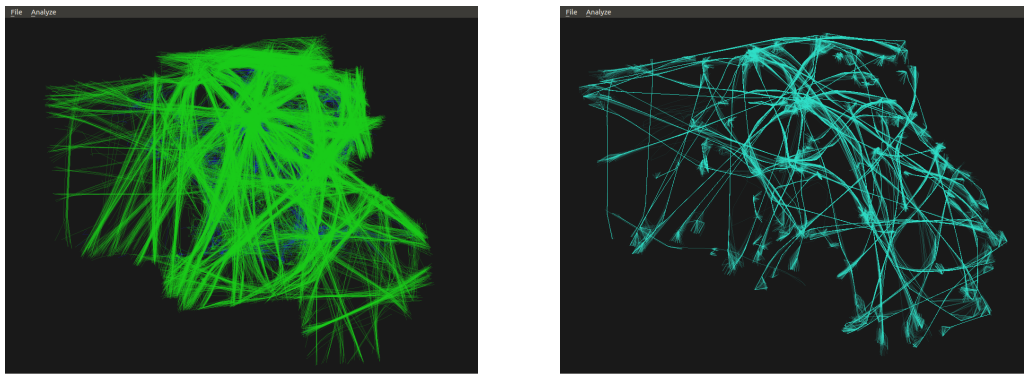


Figure 3. Initial and Bundled traffic of the 24/02/2013.

167 expected, a route-like network emerges. In such a case, since the traffic comes from an already
 168 organized situation, the recovered network is indeed a subset of the route network in the french
 169 airspace. Please note that there is a trade-off between density concentration and minimal curvature
 170 of the recovered trajectories, as already mentioned. The kernel bandwidth was chosen empirically in
 171 the example presented, with the aid of visual interaction.

172 In the second example of Figure 4, the problem of automatic conflict solving is addressed. In the
 173 initial situation, aircraft are converging to a single point, which is unsafe. Air traffic controllers will
 174 proceed in such a case by diverting aircraft from their initial flight path so as to avoid each other, but
 175 only using very simple maneuvers. An automated tool will make a full use of the available airspace
 176 and the resulting set of trajectories may fail to be manageable by a human: in the event of a system
 177 failure, no backup can be provided by controllers. The entropy minimization procedure was added to
 178 an automated conflict solver in order to end up with flight paths still tractable by humans. The final
 179 result is shown on the right part of Figure 4, where encounters no longer exists but aircraft are bound
 180 to simple trajectories, with a merging and a splitting point. Note that since the automated planner
 181 acts on velocity, all aircraft are separated in time on the inner part.

183 4. Conclusion and future work

184 Algorithms coming from the field of shape spaces emerge as a valuable tool for applications
 185 in ATM. In this work, the foundations of a post processing procedure that may be applied after an
 186 automated flight path planner are presented. Entropy minimization makes straight segments bundle

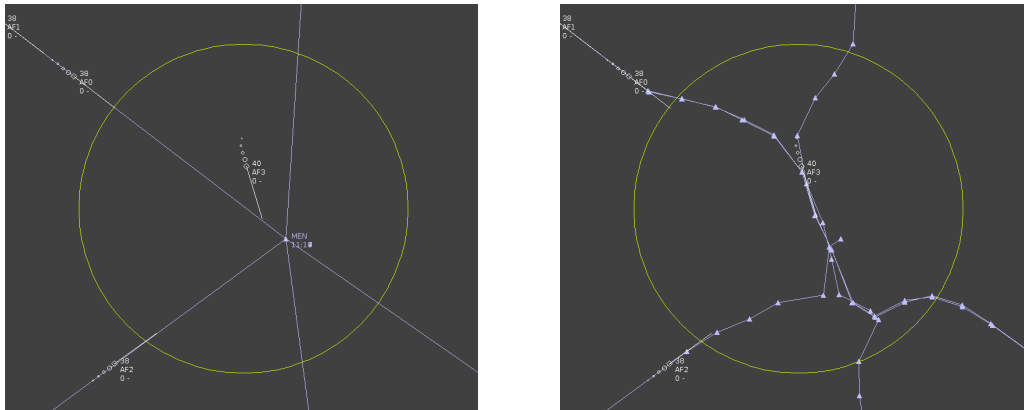


Figure 4. Initial flight plans and final ones.

187 emerge, which fulfills the operational requirements. Computational efficiency has to be improved
 188 in order to release an usable building block for future ATM systems. One way to address this
 189 issue is to compute kernel density estimators using GPUs which excel in this kind of task, very
 190 similar texture manipulations. Furthermore, statistical properties such as the optimal choice of the
 191 bandwidth parameter in the kernel estimation should be explored in more details in the next step of
 192 this work.

193 Another important point that must be addressed in future works deals with the flight paths
 194 that are very similar shape but are oriented in opposite direction. As the spatial density is not
 195 sensitive to the directional information, the entropy based procedure presented in this paper will
 196 tend to aggregate flight paths that should be sufficiently separated in order to prevent hazardous
 197 encounters. In [7], a notion of density based on position and velocity is developed. This work relies
 198 on a Lie group modeling as an unifying state representation that take into account the direction and
 199 the position of the curves. The curve system entropy has been extended to this setting.

200 Bibliography

- 201 1. L. Ambrosio, N. Gigli, and G. Savaré. *Gradient Flows: In Metric Spaces And In The Space Of Probability Measures*.
 202 Lectures in math. Birkhäuser, 2005.
- 203 2. V. A. Epanechnikov. Non-parametric estimation of a multivariate probability density. *Theory of Probability*
 204 *& Its Applications*, 14(1):153–158, 1969.
- 205 3. EUROCONTROL/NMD/STATFOR. Eurocontrol seven-year forecast. [https://](https://www.eurocontrol.int/sites/default/files/content/documents/official-documents/forecasts/seven-year-flights-service-units-forecast-2014-2020-sep2014.pdf)
 206 [www.eurocontrol.int/sites/default/files/content/documents/official-documents/forecasts/](https://www.eurocontrol.int/sites/default/files/content/documents/official-documents/forecasts/seven-year-flights-service-units-forecast-2014-2020-sep2014.pdf)
 207 [seven-year-flights-service-units-forecast-2014-2020-sep2014.pdf](https://www.eurocontrol.int/sites/default/files/content/documents/official-documents/forecasts/seven-year-flights-service-units-forecast-2014-2020-sep2014.pdf), 2014.
- 208 4. W.H. Harman. Air traffic density and distribution measurements. Lincoln Laboratory, MIT, Report ATC-80,
 209 May 1979.
- 210 5. C. Hurter, O. Ersoy, and A. Telea. Smooth bundling of large streaming and sequence graphs. In *Visualization*
 211 *Symposium (PacificVis), 2013 IEEE Pacific*, pages 41–48, Feb 2013.
- 212 6. Peter W Michor and David Mumford. Riemannian geometries on spaces of plane curves. *J. Eur. Math. Soc.*
 213 *(JEMS)*, 8:1–48, 2006.
- 214 7. Florence Nicol and Stéphane Puechmorel. Unsupervised aircraft trajectories clustering: a minimum entropy
 215 approach. In *ALLDATA 2016, The Second International Conference on Big Data, Small Data, Linked Data and Open*
 216 *Data*, Lisbon, Portugal, 2016. IARIA.
- 217 8. Emanuel Parzen. On estimation of a probability density function and mode. *Ann. Math. Statist.*,
 218 33(3):1065–1076, 09 1962.
- 219 9. Murray Rosenblatt. Remarks on some nonparametric estimates of a density function. *Ann. Math. Statist.*,
 220 27(3):832–837, 09 1956.

- 221 10. G.P. Roussos, D.V. Dimarogonas, and K.J. Kyriakopoulos. Distributed 3d navigation and collision avoidance
222 for nonholonomic aircraft-like vehicles. In *Proceedings of the 2009 European Control Conference*, 2009.
- 223 11. D.W. Scott. *Multivariate Density Estimation: Theory, Practice, and Visualization*. A Wiley-interscience
224 publication. Wiley, 1992.
- 225 12. B.W. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman & Hall/CRC Monographs on
226 Statistics & Applied Probability. Taylor & Francis, 1986.
- 227 13. W. Sun and Y.X. Yuan. *Optimization Theory and Methods: Nonlinear Programming*. Springer Optimization and
228 Its Applications. Springer US, 2006.

229 © 2016 by the authors. Submitted to *Entropy* for possible open access publication under the terms and
230 conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>)