EUROGRAPHICS 2010 / T. Akenine-Möller and M. Zwicker
(Guest Editors)

Volume 29 (2010), Number 2

Uncertain 2D Vector Field Topology

Mathias Otto [1], Tobias Germer [1], Hans-Christian Hege [2] and Holger Theisel [1]

[1]Visual Computing Group, Otto von Guericke Universität Magdeburg, Germany
[2]Visualization and Data Analysis Department, Zuse Institute Berlin, Germany

Abstract

We introduce an approach to visualize stationary 2D vector fields with global uncertainty obtained by considering the transport of local uncertainty in the flow. For this, we extend the concept of vector field topology to uncertain vector fields by considering the vector field as a density distribution function. By generalizing the concepts of stream lines and critical points we obtain a number of density fields representing an uncertain topological segmentation. Their visualization as height surfaces gives insight into both the flow behavior and its uncertainty. We present a Monte Carlo approach where we integrate probabilistic particle paths, which lead to the segmentation of topological features. Moreover, we extend our algorithms to detect saddle points and present efficient implementations. Finally, we apply our technique to a number of real and synthetic test data sets.

1. Introduction

The consideration of uncertainty is one of the most relevant problems in visualization [Joh04]. A variety of methods has been introduced to represent uncertainty in scalar, vector, and tensor fields. Most of them have in common that uncertainty is a local property of the field, i.e., it is obtained by a locally computed or measured process. Vector fields describing flow phenomena may also contain a local uncertainty which is due to noise, measurement/simulation errors, uncertain simulation parameters, initial and boundary conditions or the inherent randomness due to turbulence. However, in flow fields this uncertainty is transported along the flow, yielding a global uncertainty. Even if the integration of a particle starts in a rather certain region (i.e., in a region where it is almost sure that the velocity has a certain value), it may be transported into uncertain regions, making statements about its behavior based on a local analysis impossible. Informally spoken: if the flow goes from an area A with low local uncertainty to an area B with high local uncertainty, the global uncertainty in A is high because of the uncertainty of the outgoing flow.

In order to visualize global uncertainty in flows, we have to build upon approaches which incorporate a global analysis of the field. Therefore, we rely on topological methods. Our input are N 2D steady vector fields describing N different simulations/measurements of the same flow phenomenon. Out of them, we model the local uncertainty at any location of the domain. We describe a vector field equipped with local uncertainty as a density distribution function. In order to analyze its asymptotic flow behavior, we integrate particle density functions describing the probability that a particle is at a certain location. For this setup we define the concepts of stream lines and critical points. The resulting uncertain topological segmentation consists of a number of density distribution functions which represent the probability that a particle starting from a particular location will end in an (uncertain) source or sink. We visualize them as well as the uncertain critical points as height fields.

2. Related Work

Different types of uncertainty can appear during the whole visualization process. [GS06, JS03, PWL97] give an overview of existing uncertainty based visualization approaches. Uncertainty has been considered in fields of visualization, like isosurface [DKLP02, GR02, RLBS03, Bro04] and information visualization [SPB08].

For flow and vector data, a number of uncertainty visualization techniques have been presented. [LPSW96, WSF*95] present glyph based approaches. [SJK04] describes uncertainty patterns obtained by a reaction-diffusion model. [BWE05] introduces texture based flow visualization techniques incorporating uncertainty which is represented by cross advection and error diffusion. Furthermore, [BWE06] uses additional color schemes for representing uncertainty.

[ZDG*08] present an approach to visualize uncertainty in bidirectional vector fields. Note that all these previous approaches to uncertainty in flow visualization consider only *local* uncertainty and do not incorporate its transport. One approach to automatically analyze the reliability of flow predictions in a global context is given in [Hal02] where it is shown that under certain conditions rather unreliable input data gives still a reliable flow prediction.

Topological methods are a standard tool for vector field visualization. They gained a rather high popularity because they express even complex flow behavior by means of only a limited number of graphical primitives. We refer to [PVH*03, LHZP05] for state of the art reports in flow topology. One step towards the consideration of uncertainty in topological visualization has been taken in the context of DT-MRI data where a probabilistic fiber tracking is applied [MIW*07, STS07].

3. Vector Fields with Uncertainty

In this paper we consider stationary 2D vector fields. In classical vector field visualization such a vector field is described as a map from a 2D domain $\mathbf{D}$ into $\mathbb{R}^2$ as

$$\mathbf{v}_c(x,y) = \begin{pmatrix} u_c(x,y) \\ v_c(x,y) \end{pmatrix} \quad , \quad (x,y) \in \mathbf{D}. \tag{1}$$

In the following, we call this a *certain vector field*. If uncertainty comes into play, there is no unique vector assigned to a point (x,y) but rather a probability distribution of vectors, leading to the following definition:

Definition 1 A stationary 2D *uncertain vector field* over the domain $\mathbf{D}$ is a 4D scalar field $\rho_v(x,y;u,v)$ with

- $(x,y) \in \mathbf{D}$ and $(u,v) \in \mathbb{R}^2$
- $\rho_v(x,y;u,v) \geq 0 \quad (x,y) \in \mathbf{D}$ and $(u,v) \in \mathbb{R}^2$
- $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \rho_v(x,y;u,v)\, du\, dv = 1 \quad$ for all $\quad (x,y) \in \mathbf{D}$.

The value $\rho_v(x,y;u,v)\, du\, dv$ denotes the probability that at the location (x,y) the vector field has some value in the range $[u,u+du] \times [v,v+dv]$.

3.1. Examples of Uncertain Vector Fields

A certain vector field as described in (1) is a special case of an uncertain vector field and described by the 2D Dirac delta distribution $\rho_v(x,y;u,v) = \delta(u-u_c(x,y)\,,\, v-v_c(x,y))$. A *Gaussian distribution 2D vector field* is given by a 2D vector field $\mathbf{v}_c(x,y)$ and a 2×2 symmetric positive definite second order tensor field $\mathbf{T}(x,y)$. For a location (x,y) it assumes a Gaussian distribution of the vectors around the mean $\mathbf{v}_c$ with the standard deviation encoded in the covariance matrix $\mathbf{T}$:

$$\rho_v(x,y;u,v) = \frac{1}{2\pi\sqrt{\det(\mathbf{T})}}\; e^{-\frac{1}{2}(\mathbf{v}-\mathbf{v}_c)^T\,\mathbf{T}^{-1}\,(\mathbf{v}-\mathbf{v}_c)} \tag{2}$$

with $\mathbf{T} = \mathbf{T}(x,y)$, $\mathbf{v} = (u,v)^T$. Fig. 1 gives an illustration.

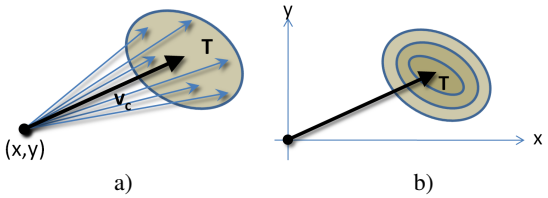

Figure 1: *Gaussian distribution vector field; a) at a location (x,y) the distribution is defined by a vector $\mathbf{v}_c$ and a 2D tensor $\mathbf{T}$ which is visualized as an ellipse; b) isolines of the corresponding distribution field ρ_v.*

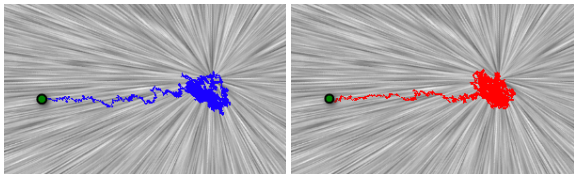

Figure 2: *Integration of 2 particles in the Gaussian distribution field (3) starting from the same location.*

3.2. Integrating Stream Lines

Classical stream line integration is not defined for uncertain vector fields as there is no unique vector at a location (x,y). To illustrate this, we consider a Gaussian distribution field (2) with

$$\mathbf{v}_c(x,y) = -0.4\begin{pmatrix} x \\ y \end{pmatrix} \quad , \quad \mathbf{T}(x,y) = \begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix} \tag{3}$$

and start an "uncertain Euler integration" of a particle at $(x,y) = (-9.5,0)$ in the following way: at the position (x_i,y_i), a vector $\mathbf{v}(x_i,y_i)$ is randomly chosen by evaluating the distribution function $\rho_v(x_i,y_i;u,v)$. Fig. 2 shows the integration of 2 different particles. Note that the particles follow different paths but have a similar global behavior: they first move towards the point $(0,0)$, then (due to the uncertainty) move randomly around it without leaving its neighborhood. This behavior does not depend on the step size of the integration: changing the step size effects the shapes and complexity of the curves but not their global behavior. Also note that the particles produce an ellipse shaped path pattern corresponding to $\mathbf{T}$.

This example shows that for uncertain vector fields we have to advect *particle density functions* instead of particular particles.

Definition 2 A *particle density function* over the domain $\mathbf{D}$ is a 2D time–dependent scalar field $p(x,y;t)$ with $(x,y) \in \mathbf{D}$, $t \in \mathbb{R}^+$ and

- $p(x,y;t) \geq 0$ for all $(x,y) \in \mathbf{D}$ and $t \geq 0$

- $\int\int_{\mathbf{D}} p(x,y;t)\,dx\,dy \leq 1$ for all $t \geq 0$.
  (We use $\leq$ instead of $=$ because particles may leave the domain during integration.)

The value $p(x,y;t)\,dxdy$ denotes the ratio of particles in $[x,x+dx]\times[y,y+dy]$ in relation to the initial number of particles in $\mathbf{D}$ at $t=0$.

Now we consider the spatial transport of particle densities by an uncertain vector field ρ_v. In order to avoid boundary effects we consider the infinite domain $\mathbf{D}=\mathbb{R}^2$. The densities are represented by virtual particles without inertia that are transported by the uncertain vector field ρ_v. Let Δt be a time interval that is short enough that the virtual particles can be considered to move on a straight line, i.e., in a constant vector field, during this interval. Then at time $t+\Delta t$ the number of particles in an infinitesimal volume $dxdy$ at some location (x,y) is the sum of the numbers of particles in cells $dr\,ds$ at all locations (r,s) times the probabilities that they are transported from (r,s) to (x,y) in time Δt, i.e., they experience a velocity $((x-r)/\Delta t,(y-s)/\Delta t)^T$. These probabilities are given by $\rho_v(r,s;\frac{x-r}{\Delta t},\frac{y-s}{\Delta t})\,d(\frac{x-r}{\Delta t})\,d(\frac{y-s}{\Delta t})$. After division by the cell volumes we have $dxdy=drds$. This yields the following expression for the transport of particle densities:

$$\begin{aligned} & p(x,y;t+\Delta t) \\ = & \int\int_{\mathbf{D}} p(r,s;t)\,\rho_v(r,s;\frac{x-r}{\Delta t},\frac{y-s}{\Delta t})\,d(\frac{x-r}{\Delta t})\,d(\frac{y-s}{\Delta t}) \\ = & \frac{1}{\Delta t^2}\int\int_{\mathbf{D}} p(r,s;t)\,\rho_v(r,s;\frac{x-r}{\Delta t},\frac{y-s}{\Delta t})\,dr\,ds \end{aligned} \tag{4}$$

A formal definition of a stream line in an uncertain vector field is given by:

Definition 3 Given an uncertain vector field $\rho_v(x,y;u,v)$, a *stream line* starting at the particle density function $p_0(x,y)$ is a time-dependent particle density function $p(x,y;t)$ with

$$\begin{aligned} p(x,y;t_0) &= p_0(x,y) \\ \frac{\partial p(x,y;t)}{\partial t} &= \lim_{\Delta t\to 0}\frac{p(x,y;t+\Delta t)-p(x,y;t)}{\Delta t} \\ p(x,y;t+\Delta t) &= (4). \end{aligned}$$

Note that definition (3) defines a forward integration of p in ρ_v. A backward integration is obtained by a forward integration of p in $\rho_v(x,y;-u,-v)$. Definition (3) also leads to uniqueness of streamlines that are started at the same particle densitiy function $p_0(x,y)$. However, $\frac{\partial p}{\partial t}$ can generally be obtained only by numerical differentiation. As an example for particle density integration, imagine that at a time t_i all particles are at a location $(r,s)\in\mathbf{D}$, i.e., that $p(x,y;t_i)=\delta(x-r,y-s)$. Applying one step of an Euler integration with the step size Δt gives the new particle distribution function $p(x,y;t_{i+1})$ at the time $t_{i+1}=t_i+\Delta t$ as

$$p(x,y;t_{i+1}) = \frac{1}{\Delta t^2}\,\rho_v\left(r,s;\frac{x-r}{\Delta t},\frac{y-s}{\Delta t}\right).$$

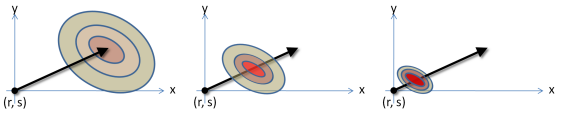

Figure 3: *One Euler integration step of a particle density function initially concentrated at (r,s) with different step sizes Δt: a) $\Delta t=1$, b) $\Delta t=0.5$, c) $\Delta t=0.25$; shown are the new particle density functions $p(x,y;t_i+\Delta t)$ as isolines.*

Note that definition (2) guarantees that $\int\int_{\mathbf{D}} p(x,y;t_{i+1})\,dxdy = 1$ for any positive Δt if $\mathbf{D}=\mathbb{R}$. Fig. 3 illustrates one step of this Euler integration if ρ_v is a Gaussian distribution field: the smaller Δt, the closer the peak of $p(x,y;t_{i+1})$ comes to (r,s) and the higher this peak is.

3.3. Critical Point Distributions

Once the concept of stream lines is established, we can make a topological analysis of an uncertain vector field: starting from every point $(r,s)\in\mathbf{D}$ (i.e., from $p(x,y;t_0)=\delta(x-r,y-s)$) we start a stream line integration and observe its asymptotic behavior for $t\to\infty$ which generally converges to a *critical point distribution*:

Definition 4 The particle density function $p_0(x,y)$ is a *critical point distribution* of ρ_v if for a stream line integration starting at $p(x,y;t_0)$ it holds $\frac{\partial p}{\partial t}=0$.

In the following we use the terms critical point distribution and critical point synonymously whenever the context is clear.

Lemma 1 Given are n critical points $p_1(x,y),\ldots,p_n(x,y)$ of ρ_v. Then any positive linear combination of critical points is a critical point as well: $\sum_{i=1}^n \alpha_i p_i$ is a critical point for any $0\leq\alpha_1,\ldots,\alpha_n\leq 1$ and $\sum_{i=1}^n \alpha_i \leq 1$.

Note that the linear combination is usually not convex, i.e., not $\sum_i \alpha_i = 1$, as part of the flow leaves the domain. Lemma 1 states that critical points of uncertain vector fields are not isolated but building a continuum of critical points. The proof follows directly from definition 3. In order to make a topological analysis, we have to find a finite number of linearly independent critical points such that every critical point can be represented as a positive linear combination of them.

Definition 5 A critical point $p(x,y)$ is a *sink distribution* (or a stable critical point) of ρ_v if the forward integration of any small perturbation of p converges to p. The critical point $p(x,y)$ is a *source distribution* of ρ_v if the backward integration of any small perturbation of p converges to p. The critical point $p(x,y)$ is a *saddle distribution* of ρ_v if both forward and backward integration are unstable, i.e., diverge from p for a small perturbation of p.

Note that sources and sinks play a different role than saddles for the definition of the topological skeleton of (certain) vector fields: sources and sinks define the number of different areas to be segmented (every segmented area corresponds to a pair of a source and a sink), while saddles are the starting points of the separating stream lines. Similar to this, we also focus on sources and sinks to define the segmentation of uncertain fields.

Definition 6 Given an uncertain vector field $\rho_v(x,y;u,v)$, a sequence of sinks $(p_1(x,y),...,p_n(x,y))$ is called *spanning sink sequence* if $p_1,...,p_n$ are linearly independent and every sink $p(x,y)$ of ρ_v can be uniquely described as $p=\sum_{i=1}^{n}\alpha_i\,p_i$ with $0\leq\alpha_1,...,\alpha_n\leq 1$ and $\sum_{i=1}^{n}\alpha_i\leq 1$. Similarly, a sequence of sources $(\tilde{p_1}(x,y),...,\tilde{p_m}(x,y))$ is called *spanning source sequence* if $\tilde{p_1},...,\tilde{p_m}$ are linearly independent and every source $\tilde{p}(x,y)$ of ρ_v can be uniquely described as $\tilde{p}=\sum_{i=1}^{m}\beta_i\,\tilde{p_i}$ with $0\leq\beta_1,...,\beta_m\leq 1$ and $\sum_{i=1}^{m}\beta_i\leq 1$.

Fig. 4a gives an illustration: the Gaussian distribution field (2) with

$$\begin{aligned}\mathbf{v}_c(x,y) &= \begin{pmatrix} -x(1-x)\,(1+x)\,(1-y^2)-xy^2 \\ y(1-y)\,(1+y)\,(1-x^2)+yx^2 \end{pmatrix}\\ \mathbf{T}(x,y) &= \begin{bmatrix} 0.25 & 0 \\ 0 & 0.25 \end{bmatrix}\end{aligned} \tag{5}$$

over the domain $[-1.5,1.5]^2$. There, the topological skeleton of $\mathbf{v}_c$ consists of 5 critical points: a saddle at $(0,0)$, two sources at $(-1,0)$ and $(1,0)$, and two sinks at $(0,-1)$ and $(0,1)$ (see the underlying LIC image in Fig. 4a). For the uncertain case, the spanning sink sequence consists of two sinks p_1,p_2 which are shown as blue height fields in Fig. 4a. The spanning source sequence consists of two sources $\tilde{p_1},\tilde{p_2}$, shown as red height fields in Fig. 4. The saddle $\hat{p_1}$ is represented as a yellow height surface. Note that $p_1,p_2,\tilde{p_1},\tilde{p_2},\hat{p_1}$ are almost zero in most regions except rather narrow peak regions. We draw their height fields only if they exceed a certain $\varepsilon>0$. Here we have chosen $\varepsilon=0.01$. This allows to show multiple height surfaces without visual clutter. Note that there is a one-to-one relation between the critical points of $\mathbf{v}_c$ and the spanning source sequence of ρ_v, as long as $||\mathbf{T}||_F$ is small enough, i.e., ρ_v is dominated by $\mathbf{v}_c$.

Once a spanning sink sequence $(p_1,...,p_n)$ is found, every sink p can be described as $p=\sum_{i=1}^{n}\alpha_i\,p_i$. We call $(\alpha_1,...,\alpha_n)$ the *coordinates of* p with respect to the spanning sink sequence $(p_1,...,p_n)$. In a similar way we define $(\beta_1,...,\beta_m)$ as the coordinates of a source $\tilde{p}$ with respect to the spanning source sequence $(\tilde{p_1},...,\tilde{p_m})$.

3.4. Topological Skeleton

Now we can describe our algorithm to extract the uncertain topological skeleton in the following way:
Given an uncertain 2D vector field $\rho_v(x,y;u,v)$ over the domain $\mathbf{D}$:

1. Find a spanning sink sequence $(p_1,...,p_n)$ of ρ_v.
2. Find a spanning source sequence $(\tilde{p_1},...,\tilde{p_m})$ of ρ_v.
3. For every location $(r,s)\in\mathbf{D}$:
   a. start a forward stream line integration at $p_0(r,s)=\delta(x-r,y-s)$ until it converges to the sink $p(r,s)$ or completely leaves $\mathbf{D}$
   b. start a backward stream line integration at $p_0(r,s)=\delta(x-r,y-s)$ until it converges to the source $\tilde{p}(r,s)$ or completely leaves $\mathbf{D}$
   c. compute the coordinates $(\alpha_1(r,s),...,\alpha_n(r,s))$ of $p(r,s)$ with respect to $(p_1,...,p_n)$
   d. compute the coordinates $(\beta_1(r,s),...,\beta_m(r,s))$ of $\tilde{p}(r,s)$ with respect to $(\tilde{p_1},...,\tilde{p_m})$

The result of the topological segmentation are $n+m$ scalar fields $(\alpha_1(r,s),...,\alpha_n(r,s))$ and $(\beta_1(r,s),...,\beta_m(r,s))$ with $(r,s)\in\mathbf{D}$. Informally spoken, $\alpha_i(r,s)$ describes the probability that a particle started to move at (r,s) will end in the i-th (uncertain) sink under forward integration. Similarly $\beta_j(r,s)$ gives the probability for a particle starting at (r,s) to converge to the j-th (uncertain) source under backward integration.

3.5. Saddle Points

For a complete topological analysis of the vector field, we have to detect saddle points in addition to the topological skeleton described in section 3.4. However, for uncertain vector fields finding saddle points is difficult, because they are unstable under forward and backward integration (see definition 5). To still find such points, we assume that the velocity of the (uncertain) vector field drops around critical points. To find local velocity minima, we derive an uncertain vector field ρ_g that corresponds to the gradient of the squared velocity in ρ_v. Given a sample of $\mathbf{v}_s(x,y)=(u_s,v_s)^T$ of $\rho_v(x,y;u,v)$ at a location (x,y), we can compute the squared velocity gradient $\mathbf{v}_g$ as

$$\mathbf{v}_g = \nabla(\mathbf{v}_s^2)=(J(\mathbf{v}_s))^T\mathbf{v}_s, \tag{6}$$

where $J(\mathbf{v}_s)$ is the Jacobian of $\mathbf{v}_s$. By creating many samples $\mathbf{v}_g=(J(\mathbf{v}_s))^T\mathbf{v}_s$, we can approximate the uncertain gradient field ρ_g. The uncertain sources of ρ_g correspond to local velocity minima. However, not all minima correspond to saddles. Therefore, we classify the sources of $\mathbf{v}_s$ by computing its (uncertain) Poincare-Hopf index at ρ_v. See section 4.4 for details.

3.6. Visualization

For an appropriate visualization of our topological segmentation, we want to represent both, the spanning critical point sets and the topological characterizations in a single visualization. This is possible because the critical points in the spanning critical point set are usually rather separated. Since the probability that a particle started at (r,s) moves into the $i-$th sink under forward and into the $j-$th

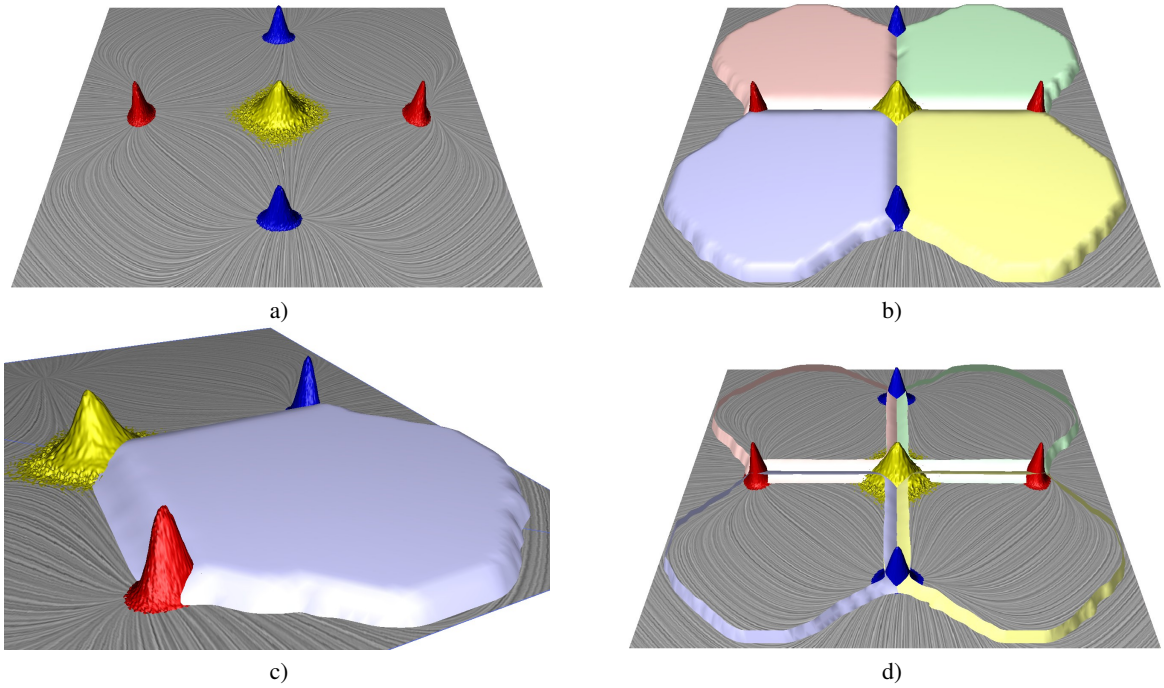

Figure 4: *Gaussian distributed vector field (6): a) LIC of v_c and uncertain critical points, b) complete segmentation as height fields, c) closeup of one sector depicting the flow from a source (red) to a sink (blue), d) skeleton reduced to uncertain areas.*

source under backward integration is described by the product $\alpha_i(r,s) \cdot \beta_j(r,s)$, we represent the $n \times m$ functions $\alpha_i\beta_j$ as randomly colored height fields but cut off areas where $\alpha_i < \varepsilon$ for a small user-chosen $\varepsilon > 0$ ($\varepsilon = 0.01$ in this paper). We use height fields, because more than two $\alpha_i\beta_j$ functions could overlap. Therefore, other techniques like color maps do not work well for this purpose. Every height surface can be interactively selected/deselected. In addition, we show the critical points as height surfaces as well: sinks in blue, sources in red, and saddles in yellow, all with the same height for visual clarity. Fig. 4b-d give an illustration for the vector field (5). Since here we have two sources and two sinks, 4 topological regions are represented as height fields (Fig. 4b). Fig. 4c shows a closeup of a region describing the flow from a source (red) to a particular sink (blue). Note that this height field is almost 1 for a rather large area, indicating a rather certain flow behavior. To focus on the uncertain regions, we cut off the height surfaces in areas of high certainty, i.e., we render them only if $\varepsilon \leq \alpha_i\beta_j \leq 1-\varepsilon$. This way, only those areas are rendered where at least two height surfaces are significantly above zero, i.e., where the flow behavior is uncertain. Fig. 4d illustrates this. In this Fig. (as well as in the following visualizations) we use a LIC (line integral convolution) of the average field $\mathbf{v}_c$ as context information for our uncertain visualization. Note that there is no direct correlation between the isolines of our height fields and the stream lines of $\mathbf{v}_c$.

4. Technical Realization

4.1. Obtaining Uncertain Vector Fields

We determine uncertainty from given data similarly to the approach proposed in [BWE05] that measures a flow phenomenon multiple times. This way our input are N vector fields over the same domain which vary due to local uncertainty. For these N samples we determine parameters of a bivariate Gaussian distribution, i.e., we obtain a mean vector field $\mathbf{v}_c$ and a tensor field of covariance matrices $\mathbf{T}$ and can hence model ρ_v. The choice of the Gaussian distribution is the most common empirical distribution of errors. It provides a fair balance of good approximation properties and efficient numerical computation. Contrary to [BWE05], we model the local uncertainty not only orthogonal to the flow direction but in every direction. This ensures stability in areas of slow flow.

4.2. Numerical Stream Line Integration

To integrate stream lines, we chose a Lagrangian Monte-Carlo approach based on probabilistic particle movements. The initial density p_0 is sampled by a high number of particles reflecting the initial distribution. Each particle is integrated by an "uncertain" Euler method: sampling the vector field $(u,v)^T$ at a certain location (x_i,y_i) yields a *random* vector respecting the distribution $\rho_v(x_i,y_i;u,v)$. In practice, we

use the Box-Muller method to sample the Gaussian distribution, which limits the sampling region to a finite radius. This sample is multiplied with the time step and added to the particle position to update its state. After a sufficient number of time steps the final distribution is approximated from the discrete spatial distribution of particles. In practice, we monitor changes in the particle density distribution over time and stop the integration if no significant changes are observed. We implement this monitoring by dividing the domain into uniform cells where the particles are counted. In our examples, the cell grid was ten times the size of the input vector field. To check for invariant distributions, we calculate the maximal difference of particle numbers within the cells every 50 time steps. We consider the distribution as stationary if this difference drops below a threshold that depends on the total number of particles and the cell size.

Special care must be taken if the particles leave the domain. At no-slip boundaries (e. g., the step in Fig. 7) , we project the particles back into the domain. If the boundary contains inflow or outflow regions, we simply stop the particles and accumulate them. This way, inflow/outflow areas act as topological source/sink and can therefore be considered for our segmentation.

4.3. Computing the Spanning Source/Sink Sequence

To compute the spanning sink sequence, we start with a Monte Carlo based integration where a high number of initial particles is equally distributed in **D**. Since this integration includes particles from every location, the critical point p_a to which it converges will contain parts of every element of the spanning sink sequence. To extract them from p_a, we compute all local maxima of p_a (including the borders of the domain) and assign a region to each of them by a flood fill algorithm. This way we find a critical point p_i for every local maximum of p_a. Before considering all p_i as elements of the spanning sink sequence, we have to check them for linear independence. If two or more p_i are linearly dependent, they are merged into one component of the spanning sink sequence. To get a spanning source sequence, a similar approach is done in backward direction.

In addition to a pure CPU version of the Monte Carlo simulation, we developed a GPU version, including interpolation of the distribution field, pseudo random number generation and path integration. Table 1 gives a comparison between the CPU and GPU versions. All timings in this paper were measured on an AMD Opteron 2218 and an NVIDIA GeForce 9800GTX platform with the GPU version.

4.4. Saddle Points

The extraction of saddle points starts similar to the computation of sources. The key difference is that we use the uncertain vector field ρ_g from section 3.5 to integrate the particles. We implemented two different approaches to achieve this:

	CPU	GPU
Euler integration	307ms	100ms
Explicit computation of ρ_g	701ms	-
Jacobian on demand	1006ms	114ms

Table 1: *Timings for one integration step for* 10^6 *particles.*

Explicit computation: The original field ρ_v transformed by its transposed Jacobian (see equation 6) possibly results in uncertainty that is no Gaussian distribution. Therefore, we explicitly compute ρ_g by storing a general density distribution function for each node (x,y) in the vector field on a discrete map. We build this map by randomly sampling the Jacobian and ρ_v at (x,y) a sufficient number of times. Finally, we can advect particles by sampling these maps.

Jacobian on demand: An alternative to the explicit storage of ρ_g is to transform ρ_v in the Euler integration only, when a sample is needed. Here, we compute a probabilistic sample of J^T and ρ_v for every particle at every time step, which is used to advect the particle. The benefit of this method is the reduced memory overhead. Therefore, we also included this in our GPU implementation. Table 1 gives a comparison of our implementations.

After finding the source distribution of ρ_g, we segment individual critical points as described in section 4.3. Finally, every critical point must be classified w.r.t. ρ_v to find saddle points. For this task, we find the probabilities for the Poincare-Hopf index by sampling the local region around the critical points and computing its index a sufficient number of times. We consider the critical point only as a saddle, if index -1 has the highest probability.

At this point, we can also extract sink and sources as critical points in ρ_g with index +1 at ρ_v. Therefore, the extraction of critical points in ρ_g serves as a preprocessing step of our algorithm. We use the found sinks and sources as starting points for the particle integration described in section 4.3. This resolves the shapes of the sources and sinks in ρ_v.

4.5. Computing Coordinates

For the integration starting at every location (r,s), we use the Monte Carlo integration and release a higher number of particles at (r,s) until the distribution converges to a sink $p_s(r,s)$. The coordinates of $p_s(r,s)$ are the relative number of particles that arrived in the respective sinks $(p_1,...,p_n)$.

In this algorithm, a high number of particles is released at every sampling position in the domain and integrated for a high number of steps. To speed this up, we exploit the spatial coherence of the particles. If the coordinates α_i are already known at a certain particle position, we can stop the particle, because its probabilities to reach sink i are expressed in α_i. Thus, the particle adds to $p_s(r,s)$ with these probabilities. To fully exploit the spatial coherence, we first integrate the particles that are started near the sinks and store the respective

coordinates at the starting points. We then successively proceed to the particles that are started in the neighborhood of the just completed region. These particles have a high probability to run into that region, where they can be stopped after only a few integration steps. This way, the total number of necessary integration steps decreases considerably. Depending on the input data we experienced a speed-up of 5× to 10×.

5. Results

To test our approach, we apply the extraction of uncertain topology to two synthetic, one measured, and two simulated data sets.

5.1. Synthetic Data Sets

Fig. 5a and 5b show two input vector fields which are obtained by considering the field

$$\mathbf{v} = \begin{cases} \begin{pmatrix} 1+1.5x \\ x(y-0.5)(5x-1) \end{pmatrix} & \text{if } x < 0 \\ \begin{pmatrix} 1-x \\ -8xy(2y-1)(y-1) \end{pmatrix} & \text{if } x \geq 0 \end{cases} \tag{7}$$

in the domain $[-0.8, 1.2] \times [-0.2, 1.2]$, where a rather strong Gaussian noise is added to the right-hand region, e.g., the region $x > 0$. From a sufficient number of them we can reconstruct the uncertain distribution field

$$\mathbf{v}_c = \mathbf{v} \ , \ \mathbf{T} = \begin{cases} \mathbf{T}(x,y) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & \text{if } x < 0 \\ \mathbf{T}(x,y) = \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{4} \end{bmatrix} & \text{if } x \geq 0 \end{cases} \tag{8}$$

This means that for $x < 0$ we have a zero local uncertainty: all input fields have exactly the same value there. Fig. 5c shows our uncertain topological skeleton consisting of one source (red height field), two sinks (blue) and one saddle (yellow). They give two topological sectors where the height fields are shown only in the uncertain areas.

This example shows two properties of our approach: firstly, the right hand side of the input fields have a strong topological complexity which is mainly due to the added Gaussian noise. Our approach can deal with it by revealing the few important topological structures. Secondly, although there is no local uncertainty in the left-hand part of the flow, our algorithm finds a rather strong global uncertainty there: for rather large areas it is uncertain to which sink a particle started from there will converge. These are the areas where the two height surfaces cross and overlap.

Fig. 6 shows a synthetic field to study the impact of increasing uncertainty on the topology. We constructed an initial vector field $\mathbf{v}_c$ as gradient of a Perlin noise scalar field. As shown in Fig. 6a, its topology contains many critical points. Then we added different amounts of uncertainty, i.e., an isotropic matrix $\mathbf{T}$ with $\|\mathbf{T}\|_F = 0.2, 2.0, 5.0$, respectively.

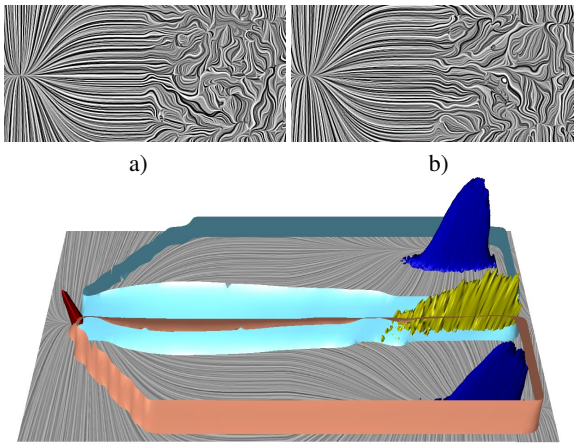

Figure 5: *Flow from a certain to an uncertain area: a, b) samples of the input fields are identical on the left-hand side but differ significantly in the right-hand region; c) the topological skeleton shows the global transport of uncertainty.*

Increasing $\|\mathbf{T}\|_F$ has two effects: the height surfaces show more overlapping areas and therefore more uncertainty, and the number of uncertain critical points decreases. Compared to Fig. 6a, the uncertain topology contains fewer structures that are more robust against variations in flow behavior.

5.2. PIV Data Set

Fig. 7 shows a PIV (particle image velocimetry) measurement of a real flow around a backward facing step. The PIV measurement was carried out 1024 times, yielding 1024 reconstructed vector fields on a regular 105x103 grid. Fig. 7a and 7b show 2 consecutive of these fields which act as input for our approach. They look rather different, indicating a higher local uncertainty. Moreover, no particular order or temporal similarity comes with the data, such that an uncertain Gaussian distribution field is the method of choice for reconstruction. Our uncertain topological visualization (Fig. 7c) consists of one saddle and two sources. In addition we interpret areas of inflow/outflow at the domain boundaries as sources/sinks as well, yielding 3 more sources and 3 more sinks (marked red and blue at the domain boundaries). The obstacle is modeled with no-slip boundaries. The visualization shows a generally high uncertainty: in rather large areas, none of the height surfaces is 1, meaning that only probabilistic statements about the flow behavior are possible. The only certain region is the upper part of the flow where the underlying LIC images are visible, indicating a rather laminar flow from the right-hand to the left-hand boundary. Fig. 7d is a close-up of Fig. 7c. The computing time for this data set was 5 minutes and 41 seconds.

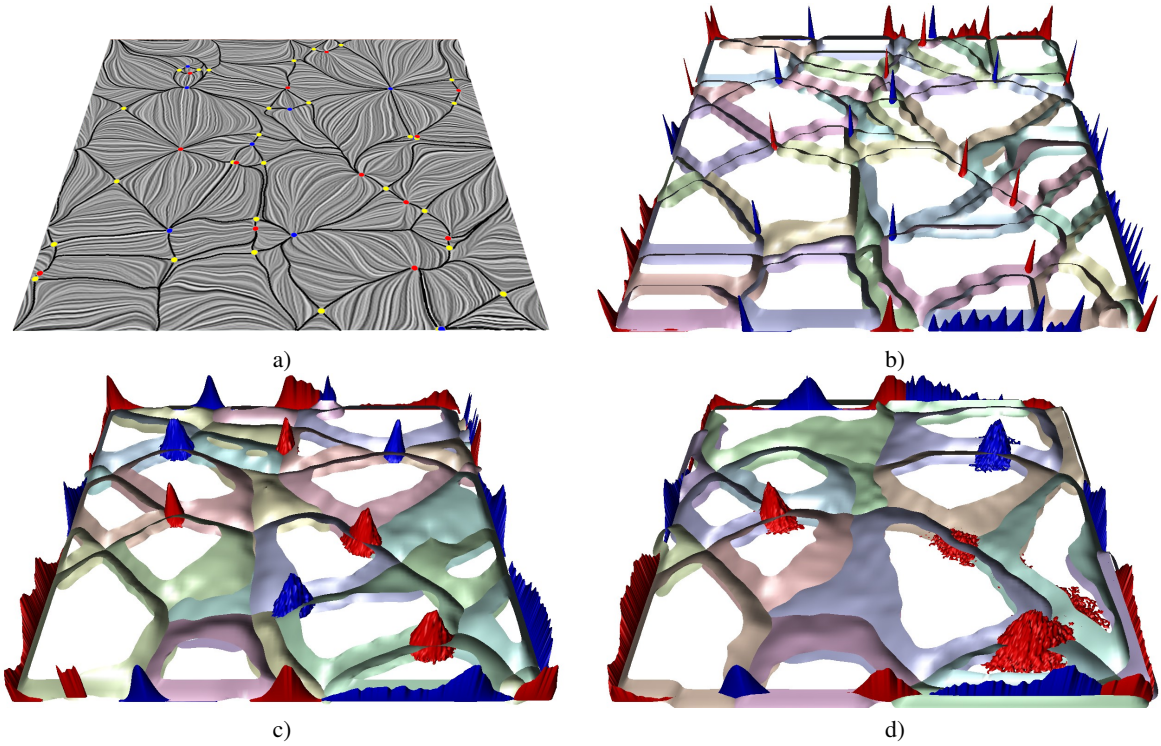

Figure 6: *Increasing the uncertainty of a random vector field: a) certain topology of mean vector field; b)* $\|\mathbf{T}\|_F = 0.2$*; c)* $\|\mathbf{T}\|_F = 2.0$*; d)* $\|\mathbf{T}\|_F = 5.0$.

5.3. Flow in the North Sea

Fig. 8 shows the visualization of simulated flow of the North Sea in the German Bight (Deutsche Bucht) between 17/10/2008 and 09/11/2008. Although the flow is 3D, it is dominated by its horizontal components and can therefore be interpreted as a 2D vector field with divergence. The flow is strongly dominated by the tides, leading to a constantly changing domain over time. For a visual analysis, we have selected all time steps of the same relative tide time (i.e., where the sea has flooded approximately the same amount of land), leading to 33 data sets over a regular 160×130 grid which describe the flow every 12 hours and 25 minutes (one tidal period). Fig. 8a and 8b show two of the input data sets. Fig. 8c shows the results of our uncertain topological analysis: including boundary inflow/outflow we have 19 sources and 21 sinks, the corresponding uncertain topology reveals both regions of certain and uncertain behavior. The computing time was 8 minutes and 56 seconds.

5.4. Flow in a Bay Area

Fig. 9 shows a data set describing (the perpendicular of) the flow of a bay area of the Baltic Sea near Greifswald (Germany). The data was given as an incomplete flow data set on a regular 115×103 grid at 25 time steps. We use these time steps as the input fields of our method. Fig. 9a-c show three of them as LIC images. Our uncertain topology revealed 74 sinks and 78 sources including regions on the boundary that where detected as critical points (see Fig. 9d). The flow falls into three classes of main behavior. In the middle part, there are large regions of a certain behavior where the underlying LIC plane is visible. Contrary, in the areas left and right of the center region strong overlap of height surfaces appears, indicating a generally more uncertain behavior. The computing time for this data set was 2 minutes and 13 seconds.

6. Discussion

Uncertain vector field topology is clearly a generalization of the classical certain vector field topology. However, even though we provide a GPU-based highly-parallel implementation, the computing time for uncertain topology is still higher than for the certain topology. This is due to the fact that we still do an expensive integration for every point, while for certain topologies only a low number of separatrices have to be integrated. These significantly high computing costs only pay off for data sets where the uncertainty is relevant and of special interest. For data sets where the uncertainty is less relevant, a faster certain topological extraction may give almost the same segmentation than the methods

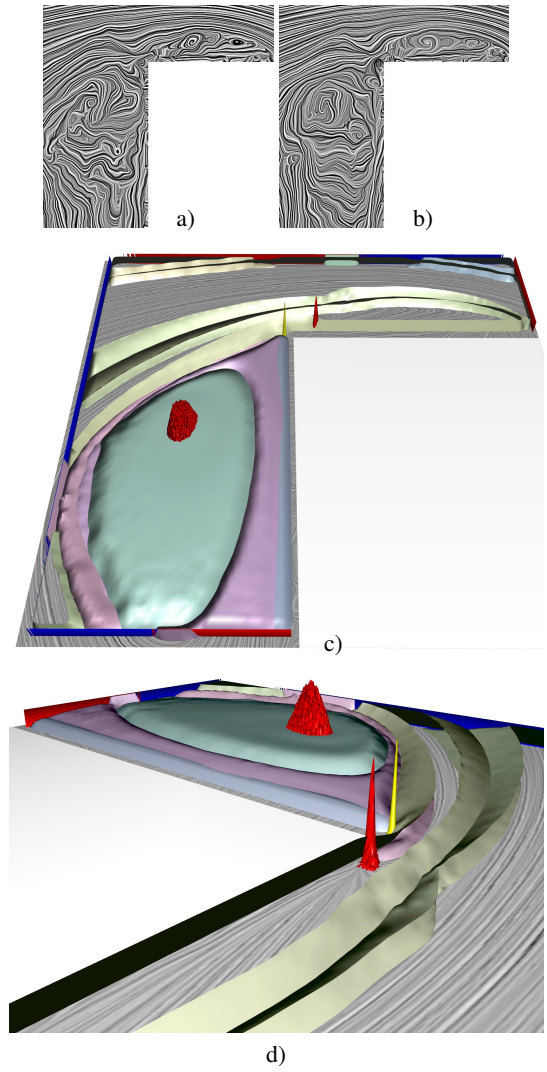

Figure 7: *PIV data set of a flow around a backward facing step: a, b) two of the 1024 data sets, c) uncertain topological skeleton, d) close-up.*

presented in this paper. Our approach is therefore not going to replace the certain topology but is a generalization particularly useful for the analysis of global uncertainty in flow fields.

Our method produces usually less uncertain critical points than the input fields have, it can therefore be interpreted as a topology simplification method. There is a number of approaches for topological simplification of (certain) vector fields [dLvL99a, dLvL99b, TSH00, TSH01]. However, these methods consider only local conditions for simplification and therefore might give different results than our global method. For example, the results in Fig. 5 are hard to achieve with local topology simplification methods.

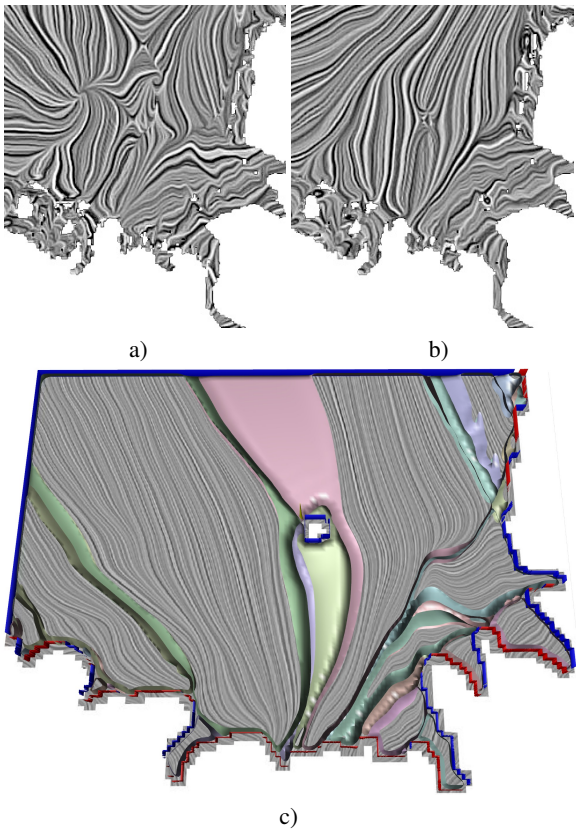

Figure 8: *Topological skeleton of a simulated flow.*

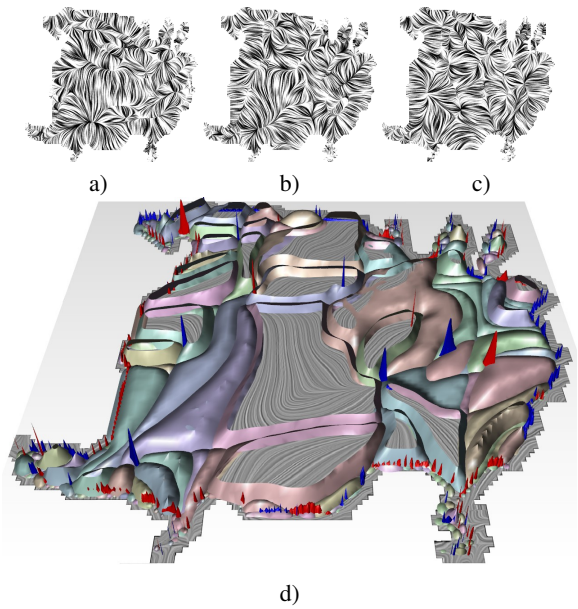

Figure 9: *Flow in a bay area: a-c) LIC of input fields; d) topological skeleton of the uncertain vector field.*

7. Conclusions

In this paper we made the following contributions:

- We have considered uncertainty in vector fields not only as a local feature but incorporated its transport as well.
- For this we have introduced uncertain vector fields as density distribution functions. Furthermore, particles and their integration are density functions as well.
- In uncertain vector fields, we have defined a topological segmentation based on an uncertain stream line integration. The resulting segmentation consists of a number of distribution functions which show the probability density that a particle starting from a given location will converge to a considered source/sink. We visualize both the critical points and the density functions as height fields.
- We have applied the technique to a number of uncertain vector field data sets which consist of different measurements or simulations of the velocity at every location of the domain.

There are a number of issues for future research. Firstly, spatial and temporal coherence of adjacent locations and particle density functions may be further exploited to speed up the approach. Secondly, the extension of the approach to 3D fields is necessary. While the extension to 3D of most concepts described in this paper is straightforward, it is practically only possible in combination with a significant acceleration of the approach. Thirdly, the extension to time-dependent fields is desired, similar to extensions to cover the behavior of stream lines and path lines in certain time-dependent fields.

8. Acknowledgements

We thank Dieter Schrader (Bundesamt fuer Seeschifffahrt und Hydrographie) for the north sea dataset, J. Boree and L. Brizzi (Lab. d'Etudes Aérodynamiques de Poitiers) for the PIV data set, the Fraunhofer IFF for providing the VDT rendering system and T. Weinkauf and J. Martinez Esturo for fruitful discussions. The work was partially funded by the German Ministry of Education and Science (BMBF) within the ViERforES project (no. 01IM08003C) and by the SemSeg project under the EU FET-Open grant 226042.

References

[Bro04] BROWN R.: Animated visual vibrations as an uncertainty visualisation technique. In *Proc. GRAPHITE* (2004), pp. 84–89.

[BWE05] BOTCHEN R. P., WEISKOPF D., ERTL T.: Texture-based visualization of uncertainty in flow fields. In *IEEE Visualization* (2005), pp. 647–654.

[BWE06] BOTCHEN R. P., WEISKOPF D., ERTL T.: Interactive Visualization of Uncertainty in Flow Fields using Texture-Based Techniques. In *Proc. Intl. Symp. on Flow Visualization* (2006).

[DKLP02] DJURCILOV S., KIM K., LERMUSIAUX P. F. J., PANG A.: Visualizing scalar volumetric data with uncertainty. *Computers and Graphics 26*, 2 (2002), 239–248.

[dLvL99a] DE LEEUW W., VAN LIERE R.: Collapsing flow topology using area metrics. In *IEEE Visualization* (1999), pp. 149–354.

[dLvL99b] DE LEEUW W., VAN LIERE R.: Visualization of global flow structures using multiple levels of topology. In *Proc. VisSym* (1999), pp. 45–52.

[GR02] GRIGORYAN G., RHEINGANS P.: Probabilistic surfaces: point based primitives to show surface uncertainty. In *IEEE Visualization* (2002), pp. 147–154.

[GS06] GRIETHE H., SCHUMANN H.: The visualization of uncertain data: Methods and problems. In *SimVis* (2006), pp. 143–156.

[Hal02] HALLER G.: Lagrangian coherent structures from approximate velocity data. *Phys. Fluids A 14* (2002), 1851–1861.

[Joh04] JOHNSON C.: Top scientific visualization research problems. *IEEE Comput. Graph. Appl. 24*, 4 (2004), 13–17.

[JS03] JOHNSON C., SANDERSON A.: A next step: Visualizing errors and uncertainty. *IEEE Comput. Graph. Appl. 23*, 5 (2003), 6–10.

[LHZP05] LARAMEE R., HAUSER H., ZHAO L., POST F.: Topology-based flow visualization, the state of the art. In *Proc. Topo-In-Vis* (2005), pp. 1–20.

[LPSW96] LODHA S. K., PANG A., SHEEHAN R. E., WITTENBRINK C. M.: Uflow: visualizing uncertainty in fluid flow. In *IEEE Visualization* (1996), p. 249.

[MIW*07] MADDAH M., III W. M. W., WARFIELD S. K., WESTIN C.-F., GRIMSON W. E. L.: Probabilistic clustering and quantitative analysis of white matter fiber tracts. In *Proc. IPMI* (2007), vol. 20, pp. 372–383.

[PVH*03] POST F., VROLIJK B., HAUSER H., LARAMEE R., DOLEISCH H.: The state of the art in flow visualization: Feature extraction and tracking. *Computer Graphics Forum 22*, 4 (2003), 775–792.

[PWL97] PANG A. T., WITTENBRINK C. M., LODHA S. K.: Approaches to uncertainty visualization. *The Visual Computer 13* (1997), 370–390.

[RLBS03] RHODES P. J., LARAMEE R. S., BERGERON R. D., SPARR T. M.: Uncertainty visualization methods in isosurface rendering. In *Eurographics* (2003), pp. 83–88.

[SJK04] SANDERSON A. R., JOHNSON C. R., KIRBY R. M.: Display of vector fields using a reaction-diffusion model. In *IEEE Visualization* (2004), pp. 115–122.

[SPB08] STREIT A., PHAM B., BROWN R.: A spreadsheet approach to facilitate visualization of uncertainty in information. *IEEE TVCG 14*, 1 (2008), 61–72.

[STS07] SCHULTZ T., THEISEL H., SEIDEL H.-P.: Topological visualization of brain diffusion MRI data. *IEEE TVCG 13*, 6 (2007), 1496–1503.

[TSH00] TRICOCHE X., SCHEUERMANN G., HAGEN H.: A topology simplification method for 2D vector fields. In *IEEE Visualization* (2000), pp. 359–366.

[TSH01] TRICOCHE X., SCHEUERMANN G., HAGEN H.: Continuous topology simplification of planar vector fields. In *IEEE Visualization* (2001), pp. 159 – 166.

[WSF*95] WITTENBRINK C. M., SAXON E., FURMAN J. J., PANG A., LODHA S.: Glyphs for visualizing uncertainty in environmental vector fields. In *IEEE TVCG* (1995), pp. 266–279.

[ZDG*08] ZUK T., DOWNTON J., GRAY D., CARPENDALE S., LIANG J.: Exploration of uncertainty in bidirectional vector fields. In *Visualization and Data Analysis* (2008), vol. 6809.