Tutorial Uncertainty and Parameter Space Analysis in Visualization

Part II Uncertainty Modeling

Hans-Christian Hege

Slides: http://www.cg.tuwien.ac.at/research/publications/2012/VisWeek-Tutorial-2012-Uncertainty
Short link: http://bit.ly/QmQfqv

Outline

1. Quantitative Representations of Uncertainty

   partially abstract, top-down

   mainly based on: *Zhenyuan Wang, George J. Klir*
   Generalized Measure Theory, Springer 2009

2. Probabilistic Modeling of Uncertain Fields

   mostly concrete, bottom-up

   mainly based on: *recent work at ZIB with*
   Kai Pöthkow, Christoph Petz, Britta Weber

Part 1: Quantitative Representations of Uncertainty

Why bother at all ?

- Consideration and quantification of uncertainties
  is of great importance in many practical applications
- Vis & VA: part of the data analysis chain + support decision taking.
- Thus we need to understand the data – including their
  shortcomings,
  value,
  relevance,
  which largely depend on presence/absence of uncertainties.

➔ We need to

- understand *quantified uncertainty* and deal with it
- perform *uncertainty quantification* by ourselves

What is Uncertainty ?

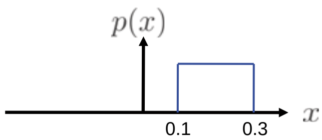
uncertainty ⇔ lack of information

- Uncertainty due to randomness — aleatoric uncertainty
  - Results by chance
  - Lack of information is objective
  - *Example*: daily quantity of rain in Seattle

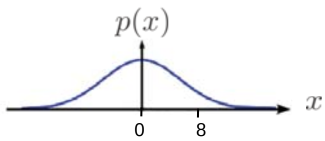
- Uncertainty due to lack of knowledge — epistemic uncertainty
  - In principle we could know, but in practice we don't know
  - Lack of knowledge is subjective
  - *Example*: birth date of last Chinese Emperor

Uncertain Propositions - Examples

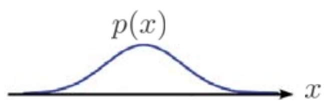
- "The value of x is between 0.1 and 0.3"



- "The value of x is normally distributed with zero mean and standard deviation 5.0"



- "The value of x is normally distributed"

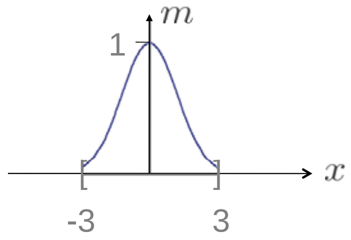


- "Bob is middle-aged"

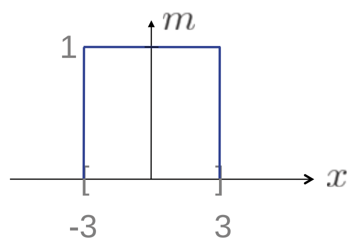
???

Insertion: Fuzzy Sets

- Let X be a nonempty set (X= the “universe” of discourse)
  Fuzzy set in X characterized by membership function $m : X \longrightarrow [0,1]$
- Example: fuzzy set $X = [-3,3] \subset \mathbb{R}$
  $m(x) = e^{-x^2}$



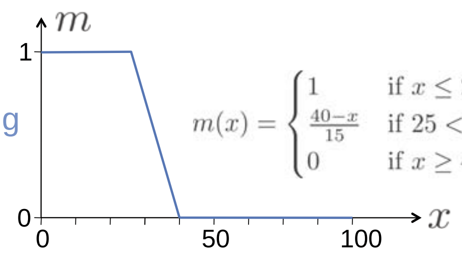
- Example: (crisp) set $X = [-3,3] \subset \mathbb{R}$
  $m(x) = 1$



Example: Modelling of Imprecise Age Statements

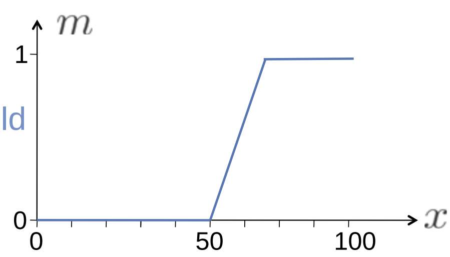
- Measure in years; age interval of human beings: $X = [0,100] \subset \mathbb{N}$

young

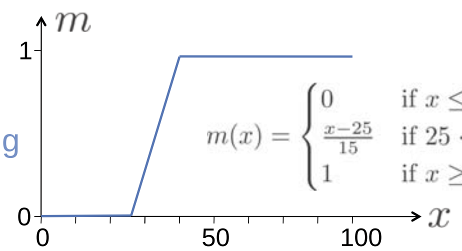


$$m(x) = \begin{cases} 1 & \text{if } x \leq 25 \\ \frac{40-x}{15} & \text{if } 25 < x < 40 \\ 0 & \text{if } x \geq 40 \end{cases}$$

old

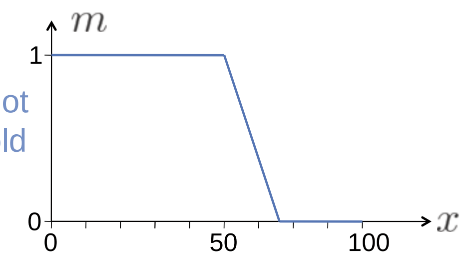


not young



$$m(x) = \begin{cases} 0 & \text{if } x \leq 25 \\ \frac{x-25}{15} & \text{if } 25 < x < 40 \\ 1 & \text{if } x \geq 40 \end{cases}$$

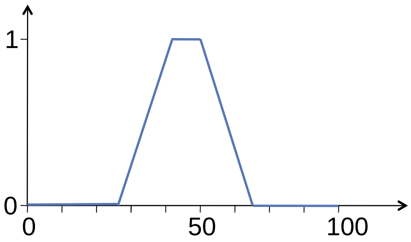
not old



Example: Modelling of Imprecise Age Statements

- Age interval of human beings: $X = [0, 100] \subset \mathbb{N}$

neither young nor old (middle-aged)



$$m(x) = \begin{cases} 0 & \text{if } x \leq 25 \\ \frac{x-25}{15} & \text{if } 25 < x < 40 \\ 1 & \text{if } x \geq 40 \\ \frac{65-x}{15} & \text{if } 50 < x < 65 \\ 0 & \text{if } x \geq 65 \end{cases}$$

Puzzles and Problems

Paradoxon of total ignorance

Is there life beyond Earth ?

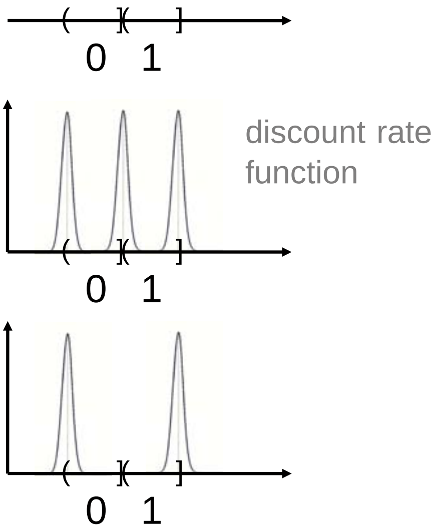
Case 1: beyond Earth:	life	no life	
Ignorant's response:	½	½	
Case 2:	animal life	plant life	no life
Ignorant's response:	1/3	1/3	1/3

Answers inconsistent: from case 2: P(life) = 2/3 > 1/2 = P(no life)
from case 1: P(animal life) = 1/4 < 1/3 = P(no life)

➔ Uniform probabilities on distinct representations of the state space are inconsistent.
➔ A probability distribution cannot model ignorance (maximal incompleteness).

Puzzles and Problems

Imprecise measurement with digital outcome :



- *Observations* at the boundaries of the intervals are *unreliable*
  ➔ they should be properly discounted

- Taking measurements for union of the 2 events
  ➔ one of the discount rate peaks is not applicable
  ➔ *same observations* produce *more evidence* for single event $0 \cup 1$ then for 2 disjoint events $0, 1$

➔ $prob(0 \cup 1) \geq prob(0) + prob(1)$ non-additive !

Mathematical Modelling of Uncertainty

- A variety of *types of uncertainty* occur in practice, including *mixtures*.
- Quantification of uncertainties, including mixtures, requires a unifying mathematical framework.
- Establishing such a mathematical framework is *difficult* ! (it already required centuries …)
- Development of such a theory is not yet fully accomplished, but silhouettes start to become visible !

What I will outline here

- What is the overall picture ?
- What are the major types of modeling ?
- What is the general mathematical framework behind ?
- Where can I find further information ?

And what not:

- **Any** technical details about the theories
- Illustrating examples

Fundamental Setting

- X : set of all elementary events (= the "universe")
- Situation with possible outcomes or occurrences of "events" $A, B, C, \ldots$
- Events $A, B, C, \ldots$ are subsets of X, i.e. elements of power set of X

➔ Tasks:

Measure the

evidence	that event A happened
degree of truth	of the statement "event A happened"
probability	that event A will happen

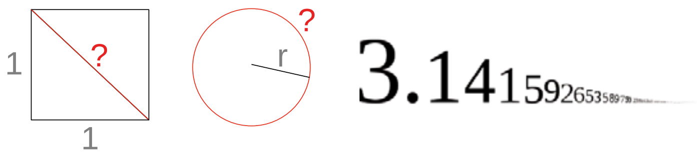
Measures in Mathematics

to measure = to assign real numbers to sets

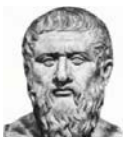
- Classical task in metric geometry: assign numbers to geometric objects for lengths, areas, or volumes
- Requirement: assigned numbers should be *invariant* under displacement of respective objects
- In ancient times: to measure = to compare with a standard unit

Measure in Mathematics

- Soon: *problem of incommensurables*



Hippasos (~500 BC)
Plato (~400 BC)



➔ Measurement is more complicated than initially thought.
It involves infinite processes and sets.

- 1854: First tool to deal with the problem: Riemann integral

➔ Enables to compute lengths, areas, volumes for complex shapes (as well as other measures).

Riemann (1826-66)

Measures in Mathematics

- ~ 1870s and 80s: Riemann integral has a number of *deficiencies*
  - Applicable only to functions with finite number of discontinuities
  - Fundamental operations of differentiation and integration are in general not reversible in the context of Riemann theory
  - Limit processes can in general not be interchanged:

$$\int_a^b \lim_{n\to\infty} f_n(x)dx \quad \text{and} \quad \lim_{n\to\infty} \int_a^b f_n(x)dx \quad \text{may differ.}$$

Measures in Mathematics

- 1898: Émile Borel developed classical measure theory

Borel (1871-1956)

  - Defined σ-algebra = class of sets that is closed under set union of countably many sets and set complement
  - Defined measure μ that associates a number $\in \mathbb{R}_0^+$ with each bounded subset in the σ-algebra
  - The measure is additive:

$$\mu(A+B) = \mu(A) + \mu(B) \qquad \text{if } A \cap B = \emptyset$$

Measures in Mathematics

- 1899-1902: Henry Lebesgue defined integral
  - Based on a measure that subsumes the Borel measure as a special case
  - Connected measures of sets and measures of functions

Lebesgue (1875-1941)

- 1933: Andrey Nikolaevich Kolmogorov developed the concept of probability measure
  - Used classical measure
  - Added: measure 1 is assigned to the universal set

Kolmogorov (1903-87)

➔ Classical Probability Theory

Measures in Mathematics

- About 50 years later: additivity requirement became a subject of controversy
  - Too restrictive to capture e.g. the *full scope of measurement*:
    - Works well under idealized error-free measurements
    - Not adequate when measurement errors are unavoidable

- The two basic types of uncertainties in relation to experiments:
  - Aleatoric: results differ each time she/he runs an experiment
    phenomenon is truly random; results „depend“ on chance
    ➔ probabilistic modeling

  - Epistemic: in principle we could know the exact results,
    but we don’t know in practice;
    due to errors that practically cannot controlled;
    ➔ non-probabilistic modeling

Measures in Mathematics

Choquet (1915-2006)

1954 Gustave Choquet developed a (potentially infinite) family of non-additive measures ("capacities")

- For each given capacity there exists a dual "alternating capacity"
- Integral based on these measures (Choquet integral)
  - non-additive
  - can be computed using Riemann or Lebesgue integration
  - applied specifically to membership functions and capacities

Dempster-Shafer Theory

Dempster (~1930 -)

Motivation: precision required in classical probability not realistic in many applications

1967 Arthur P. Dempster introduced imprecise probabilities

- Dealt with *convex sets of probability measures* rather than single measures
- For each given convex set of probability measures he introduced
  - 2 types of non-additive measures: lower & upper probabilities
    super- & supra-additive

Allows to represent probabilities imprecisely by **intervals** of real numbers.

Dempster-Shafer Theory

Shafer (~1946 –)

1976 Glenn Shafer

analyzed special types of lower & upper probabilities

called them belief & plausibility measures

- Theory based on these measures = Dempster-Shafer theory (DST) or evidence theory
- DST is capable of dealing with interval-based probabilities:
  [belief measure, plausibility measure] = ranges of admissible probabilities
- Turns out: belief measures = Choquet capacities of order ∞
  plausibility measures = alternating capacities of order ∞

Measures in Mathematics

Sugeno (1940 -)

1978 Michio Sugeno tried to compare

- membership functions of fuzzy sets
- with probabilities

not directly possible

- Generalization of additive measure analogous to generalization
  crisp sets $\longrightarrow$ *fuzzy sets*
  additive measure $\longrightarrow$ *fuzzy measure* (non-additive)
  monotone measure
- Introduced also Sugeno integral with respect to a monotone measure

Measures in Mathematics

Zadeh (1921-)

1978 Lotfi Zadeh defined:

„Possibility function“ associated with each fuzzy set (numerically: membership function)

„Possibility measure“ supremum of the possibility function in each set of concern (both for crisp and fuzzy sets)

- One of several interpretations of the “theory of graded possibilities”
- Connection to DST:
  plausibility measures = possibility measures (consonant plausibility measures)
  belief measures = necessity measures (consonant belief measures)

Classes of Uncertainty Theories

Using additive measures

- $\mu(A \cup B) = \mu(A) + \mu(B)$ expresses no interaction between events

➔ classical probability + measure theory

Using non-additive measures

- $\mu(A \cup B) > \mu(A) + \mu(B)$ expresses positive interaction between events
  ➔ synergy, cooperation, coalition, enhancement, amplification
- $\mu(A \cup B) < \mu(A) + \mu(B)$ expresses negative interaction between events
  ➔ incompatibility, rivalry, inhibition, downgrading, condensation

➔ (many) uncertainty theories + generalized measure theory

Most Utilized Uncertainty Theories + Further Reading

1. Classical Probability Theory
2. Dempster-Shafer Theory

   Simona Salicone:
   “Measurement Uncertainty: An Approach via the Mathematical Theory of Evidence”,
   Springer, 2007

   Jürg Kohlas, Paul-Andre Monney:
   “A Mathematical Theory of Hints: An Approach to the Dempster-Shafer Theory of Evidence“
   Springer, 1995
3. Possibility Theory

   Didier Dubois and Henri Prade:
   "Possibility Theory, Probability Theory and Multiple-valued Logics: A Clarification“,
   Annals of Mathematics and Artificial Intelligence 32:35-66, 2001

   Gerla Giangiacomo:
   “Fuzzy logic: Mathematical Tools for Approximate Reasoning“,
   Kluwer Academic Publishers, Dordrecht 2001

Slides: http://www.cg.tuwien.ac.at/research/publications/2012/VisWeek-Tutorial-2012-Uncertainty
Short link: http://bit.ly/QmQfqv

Thank you very much for your attention !

www.zib.de/visual

Slides: http://www.cg.tuwien.ac.at/research/publications/2012/VisWeek-Tutorial-2012-Uncertainty
Short link: http://bit.ly/QmQfqv

Slides: http://www.cg.tuwien.ac.at/research/publications/2012/VisWeek-Tutorial-2012-Uncertainty
Short link: http://bit.ly/QmQfqv

Part 2: Probabilistic Modeling of Uncertain Fields

Sources:

Probabilistic marching cubes.
Kai Pöthkow, Britta Weber, HCH
Comput. Graph. Forum 30:3, 2011 pp. 931-940.

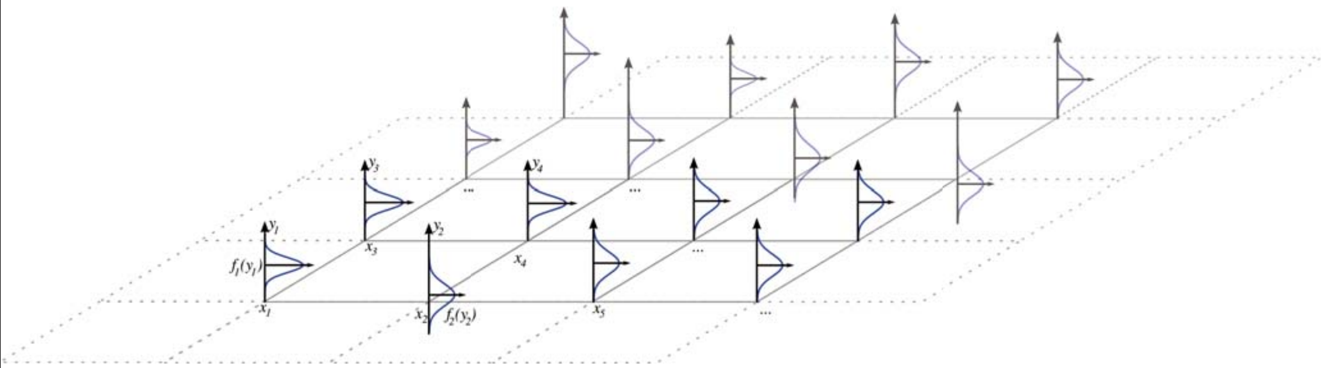
Probabilistic local features in uncertain vector fields with spatial correlation
Christoph Petz, Kai Pöthkow, HCH
Computer Graphics Forum 31:3, 2012, pp. 1325-1334.

Approximate level-crossing probabilities for interactive visualization of uncertain isocontours.
Kai Pöthkow, Christoph Petz, HCH
Int. J. Uncertainty Quantification (2012; forthcoming)

Uncertain Scalar Field

Model as *discrete random field*

For simplicity: use *Gaussian random variables*



Each field configuration: conceived as a realization of a multivariate Gaussian RV

Gaussian Random Field

Discrete random field = multivariate Gaussian RV

$$\mathbf{Y} \sim \mathcal{N}_n(\mu, \Sigma) \qquad \mu = [\mathrm{E}(Y_1), \mathrm{E}(Y_2), \ldots, \mathrm{E}(Y_n)]$$

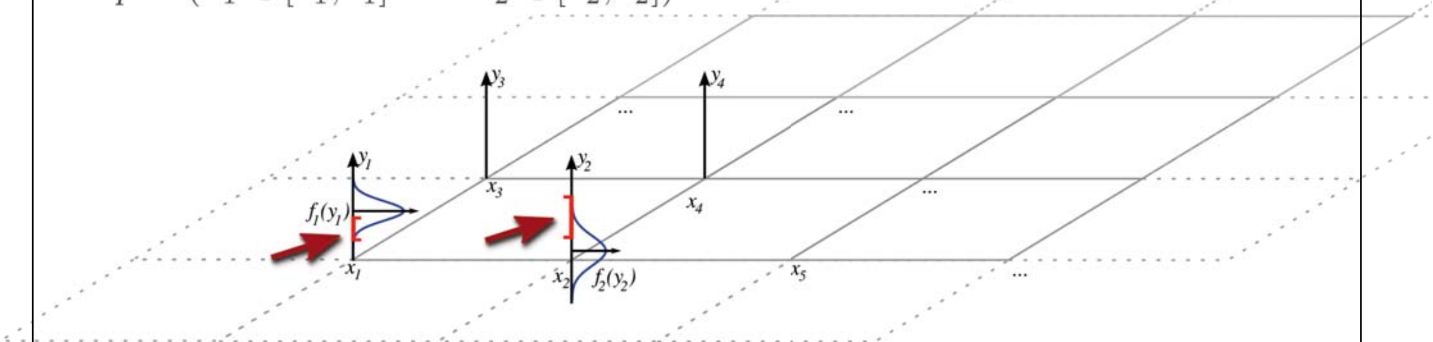
$$\Sigma = [\mathrm{Cov}(Y_i, Y_j)]_{i=1,2,\ldots,n; j=1,2,\ldots,n}.$$

$$\mathbf{Y}(\mathbf{y}) = \frac{1}{(2\pi)^{n/2} \det(\Sigma)^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{y}-\mu)^T \Sigma^{-1} (\mathbf{y}-\mu)\right)$$

Gaussian Random Field

Compute probability of locally defined events, e.g.

$prob(x_1 \in [a_1, b_1] \text{ and } x_2 \in [a_2, b_2])$



- ➔ Sum over *all* configurations that respect to predicate in the argument of $prob(\ldots)$
- ➔ Majority of variables are "integrated out" (marginalized)
- ➔ Then only a few local integrations remain

Gaussian Random Field

Marginalization:

$$\int_{-\infty}^{\infty} dy_{m+1} \cdots \int_{-\infty}^{\infty} dy_n \frac{1}{(2\pi)^{n/2} \det(\Sigma)^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{y}-\mu)^T \Sigma^{-1} (\mathbf{y}-\mu)\right)$$

$$= \frac{1}{(2\pi)^{m/2} \det(\tilde{\Sigma})^{1/2}} \exp\left(-\frac{1}{2}(\tilde{\mathbf{y}}-\tilde{\mu})^T \tilde{\Sigma}^{-1} (\tilde{\mathbf{y}}-\tilde{\mu})\right)$$

$$=: f_{\tilde{\mathbf{Y}}}(y_1, \ldots, y_m)$$

where $\tilde{\mathbf{Y}}$ is the reduced random vector and $\tilde{\mathbf{y}}$, $\tilde{\mu}$ and $\tilde{\Sigma}$ are the quantities $\mathbf{y}$, μ and Σ with $n-m$ columns/rows deleted corresponding to the marginalized variables $y_{m+1} \cdots y_n$

Gaussian Random Field

Complete random field

$$Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_m \\ \vdots \\ Y_n \end{bmatrix} \quad \mu = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_m \\ \vdots \\ \mu_n \end{bmatrix} \quad \Sigma = \begin{bmatrix} \mathrm{Cov}_{1,1} & \cdots & \mathrm{Cov}_{1,m} & \cdots & \mathrm{Cov}_{1,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathrm{Cov}_{m,1} & \cdots & \mathrm{Cov}_{m,m} & \cdots & \mathrm{Cov}_{m,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathrm{Cov}_{n,1} & \cdots & \mathrm{Cov}_{n,m} & \cdots & \mathrm{Cov}_{n,n} \end{bmatrix}$$

Local marginal distribution

$$\tilde{Y} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_m \end{bmatrix} \quad \tilde{\mu} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_m \end{bmatrix} \quad \tilde{\Sigma} = \begin{bmatrix} \mathrm{Cov}_{1,1} & \cdots & \mathrm{Cov}_{1,m} \\ \vdots & \ddots & \vdots \\ \mathrm{Cov}_{m,1} & \cdots & \mathrm{Cov}_{m,m} \end{bmatrix}$$

Probabilities of Classes of Realizations

Constrain $m \leq n$ RV Y_i to subsets S_i.

Re-order RV such that constrained ones are the first m ones.

Probability of constrained realization:

$$\mathrm{Prob}\,(Y_1 \in S_1, \ldots, Y_m \in S_m) =$$

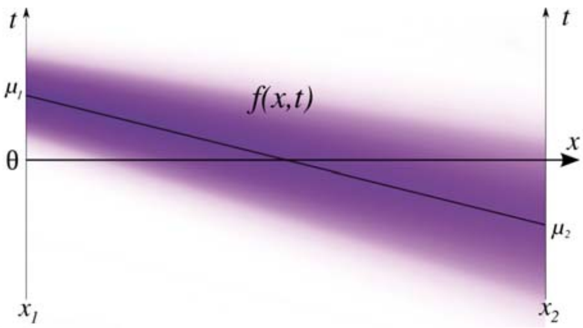
$$\int_{S_1} \mathrm{d}y_1 \ldots \int_{S_m} \mathrm{d}y_m \int_{\mathbb{R}} \mathrm{d}y_{m+1} \ldots \int_{\mathbb{R}} \mathrm{d}y_n \; f_{\mathbf{Y}}(y_1, \ldots, y_n)$$

For Gaussian distribution:

$$\int_{S_1} \mathrm{d}y_1 \ldots \int_{S_m} \mathrm{d}y_m \; f_{\tilde{\mathbf{Y}}}(y_1, \ldots, y_m)$$

Level Crossing Probabilities

For any realization (= grid function g) assume an C^0 interpolant taking its extreme values at the sample points.



Consider a particular grid cell c with vertex indices $\tilde{I} \in I$.

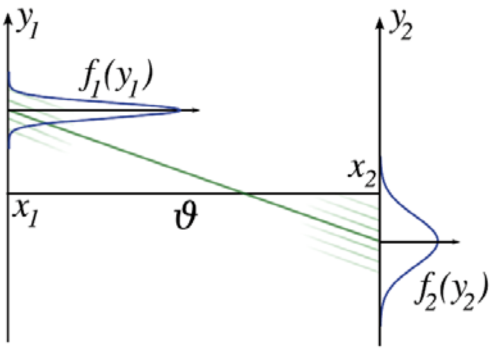
Cell c crosses ϑ-level of $g_{\{y\}}$ iff not all differences $(y_i - \vartheta)_{i \in \tilde{I}}$ have the same sign.

Level crossing probability $\mathrm{Prob}_c(\vartheta\text{-crossing})$:

Integrate $\{Y\}_{i \in \tilde{I}}$ over sets $\{y_j \in \mathbb{R} \,|\, y_j \geq \vartheta\}$ and $\{y_i \in \mathbb{R} \,|\, y_i \leq \vartheta\}$

Level Crossing Probabilities on Edges

Edge with bivariate Gaussian RV $\mathbf{Y} = [Y_1, Y_2]$



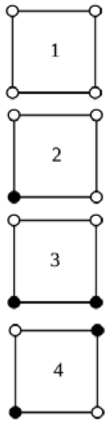
$$\mathrm{Prob}_c(\vartheta\text{-crossing}) =$$

$$= \mathrm{Prob}(Y_1 \leq \vartheta, Y_2 > \vartheta) \quad + \ \mathrm{Prob}(Y_1 > \vartheta, Y_2 \leq \vartheta)$$

$$= \int_{y_1 \leq \vartheta} \mathrm{d}y_1 \int_{y_2 > \vartheta} \mathrm{d}y_2 \ f_{\mathbf{Y}}(y_1, y_2) + \int_{y_1 > \vartheta} \mathrm{d}y_1 \int_{y_2 \leq \vartheta} \mathrm{d}y_2 \ f_{\mathbf{Y}}(y_1, y_2)$$

Level Crossing Probabilities on Faces

4 Cases (after Symmetry Reduction) $\triangleq$ Corresponding Integrals



$$P_{\vartheta,1} = \int_{(y_1 > \vartheta \wedge y_2 > \vartheta \wedge y_3 > \vartheta \wedge y_4 > \vartheta)} \mathrm{d}y_1 \int \mathrm{d}y_2 \int \mathrm{d}y_3 \int \mathrm{d}y_4 \ f_{\mathbf{Y}}(y_1, y_2, y_3, y_4)$$

$$P_{\vartheta,2} = \int_{(y_1 \leq \vartheta \wedge y_2 > \vartheta \wedge y_3 > \vartheta \wedge y_4 > \vartheta)} \mathrm{d}y_1 \int \mathrm{d}y_2 \int \mathrm{d}y_3 \int \mathrm{d}y_4 \ f_{\mathbf{Y}}(y_1, y_2, y_3, y_4)$$

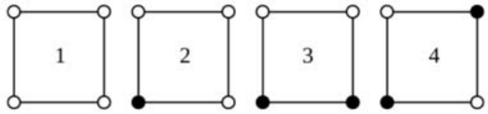
$$P_{\vartheta,3} = \int_{(y_1 \leq \vartheta \wedge y_2 \leq \vartheta \wedge y_3 > \vartheta \wedge y_4 > \vartheta)} \mathrm{d}y_1 \int \mathrm{d}y_2 \int \mathrm{d}y_3 \int \mathrm{d}y_4 \ f_{\mathbf{Y}}(y_1, y_2, y_3, y_4)$$

$$P_{\vartheta,4} = \int_{(y_1 \leq \vartheta \wedge y_2 > \vartheta \wedge y_3 \leq \vartheta \wedge y_4 > \vartheta)} \mathrm{d}y_1 \int \mathrm{d}y_2 \int \mathrm{d}y_3 \int \mathrm{d}y_4 \ f_{\mathbf{Y}}(y_1, y_2, y_3, y_4)$$

Level Crossing Probabilities on Rectangular Cells, …

Types of integrals $\triangleq$ symmetry-reduced Marching cubes cases.

In 2D: 4 distinct cases (1 non-crossing, 3 crossing)



In 3D: 15 distinct cases (1 non-crossing, 14 crossing)

In 4D: 223 distinct cases (1 non-crossing, 222 crossing)

In nD: use Polya's counting theory

Level Crossing Probabilities – Simplified

of cases (i.e. integrals) **with** level crossings grows with dimension …

Better exploit $\text{Prob}_c(\vartheta\text{-crossing}) = 1 - \text{Prob}_c(\vartheta\text{-non-crossing})$

only **2** cases **without** level crossings
➜ for all dimensions only **2** integrals!

e.g. for square cells in 2D:

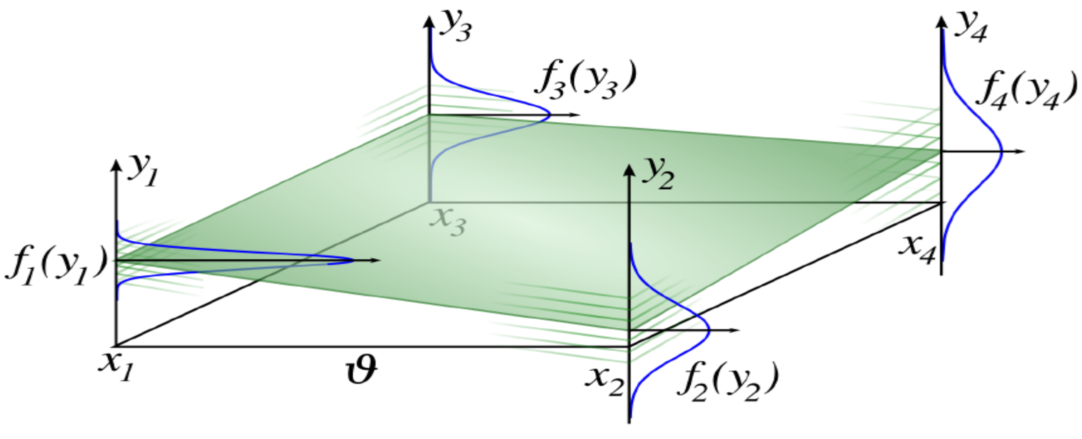
$$\text{Prob}_c(\vartheta\text{-crossing}) = 1 - \int_{\substack{(y_1 \le \vartheta \wedge y_2 \le \vartheta \wedge y_3 \le \vartheta \wedge y_4 \le \vartheta) \\ \vee (y_1 > \vartheta \wedge y_2 > \vartheta \wedge y_3 > \vartheta \wedge y_4 > \vartheta)}} \mathrm{d}y_1 \int \mathrm{d}y_2 \int \mathrm{d}y_3 \int \mathrm{d}y_4 \; f_{\mathbf{Y}}(y_1, y_2, y_3, y_4)$$

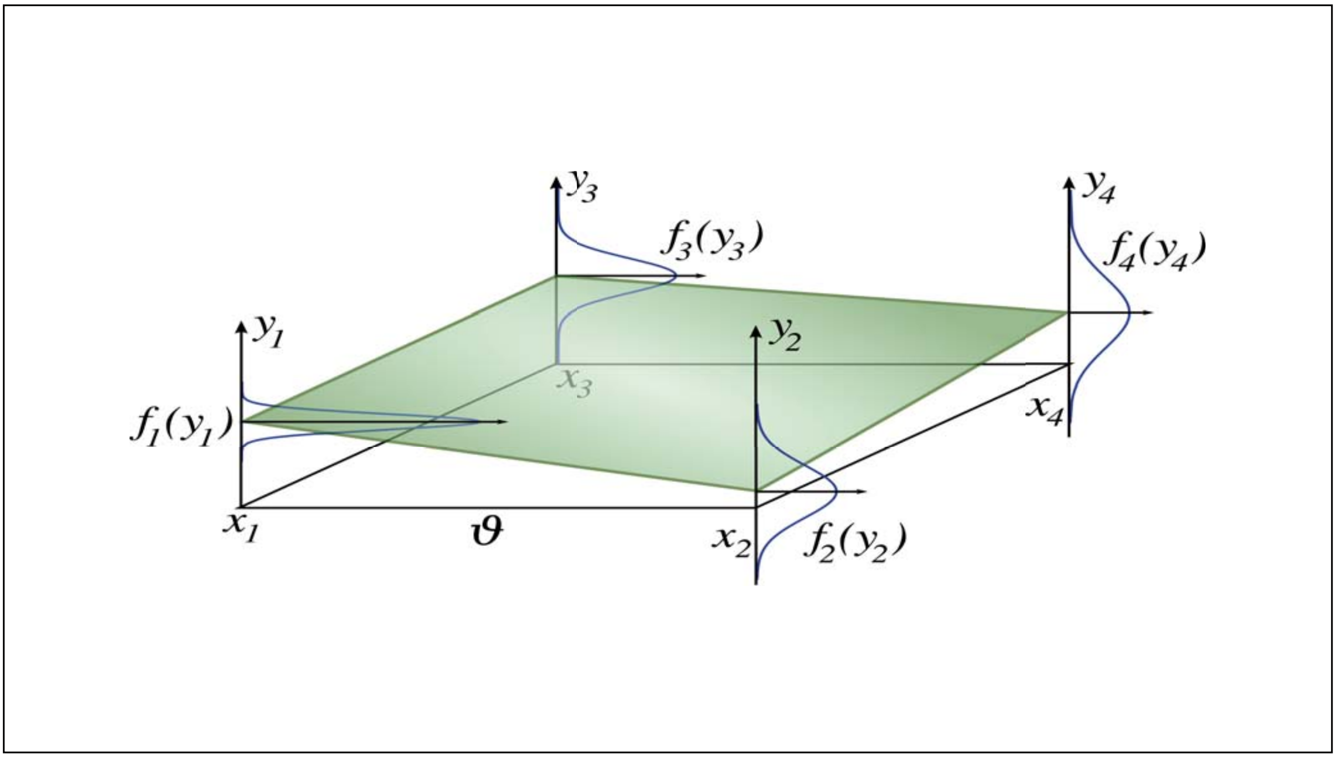
But dimension of integrals still = # vertices of geometric object !

Algorithm & Implementation

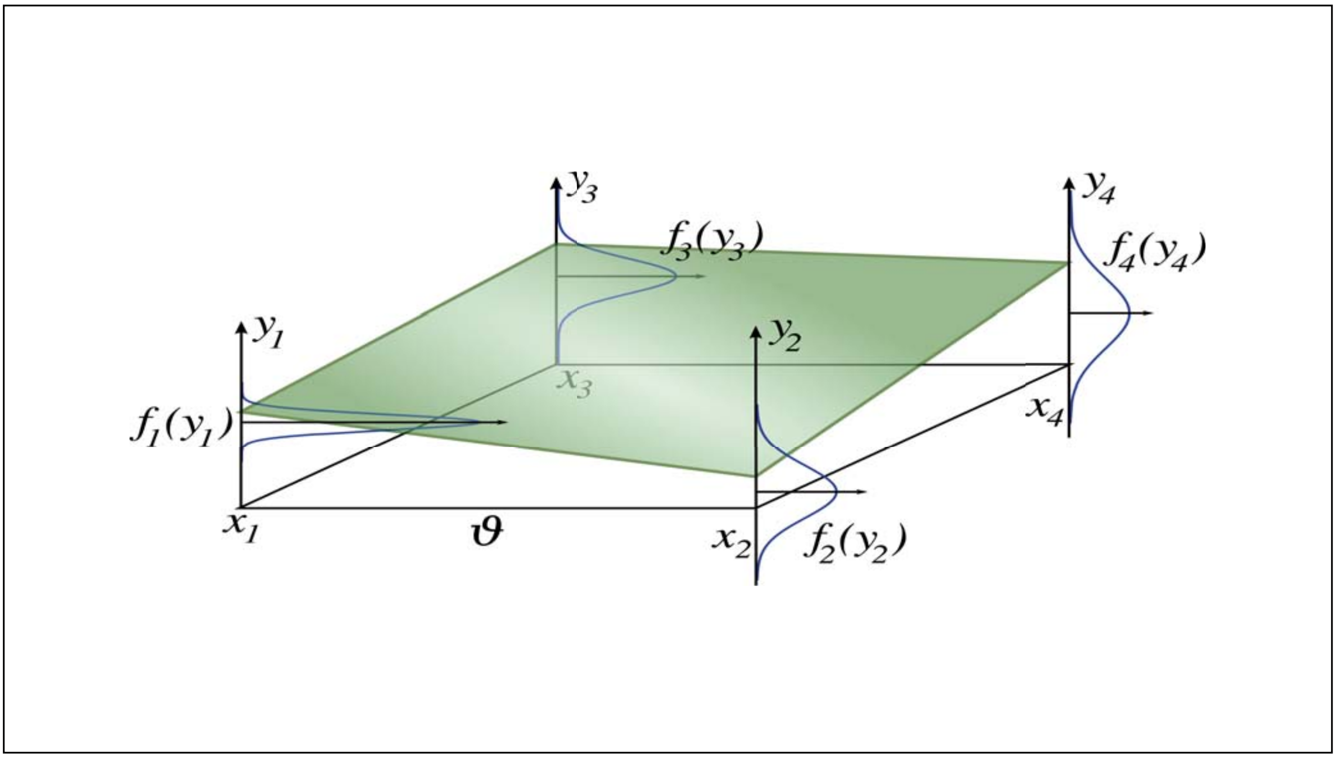
- Preprocessing
  - Estimate $\hat{\mu}_i$ for all sample points
  - Estimate $\widehat{\mathrm{Cov}}_{i,j}$ for all 2- or 3-cells
- For a given iso-value ϑ
  - Estimate crossing probabilities using Monte Carlo integration

Level Crossing Probabilities on Faces

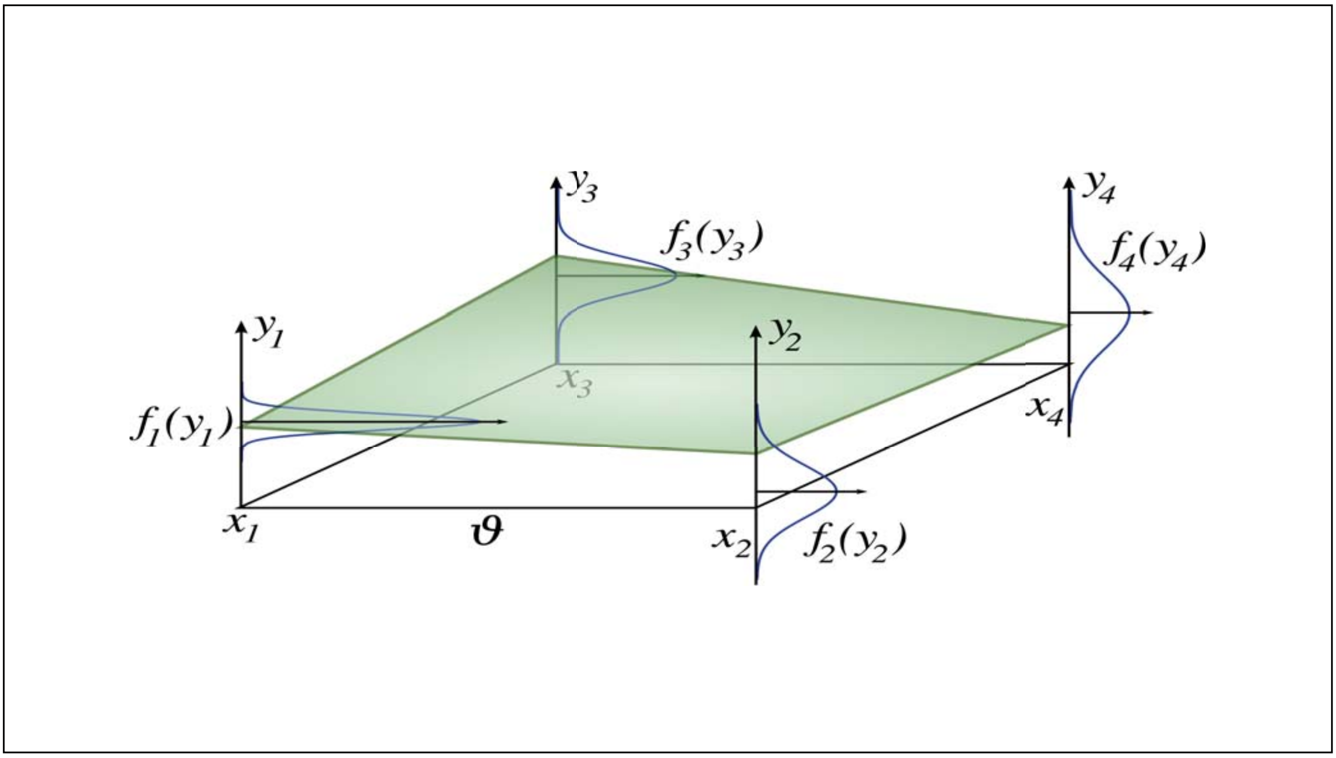




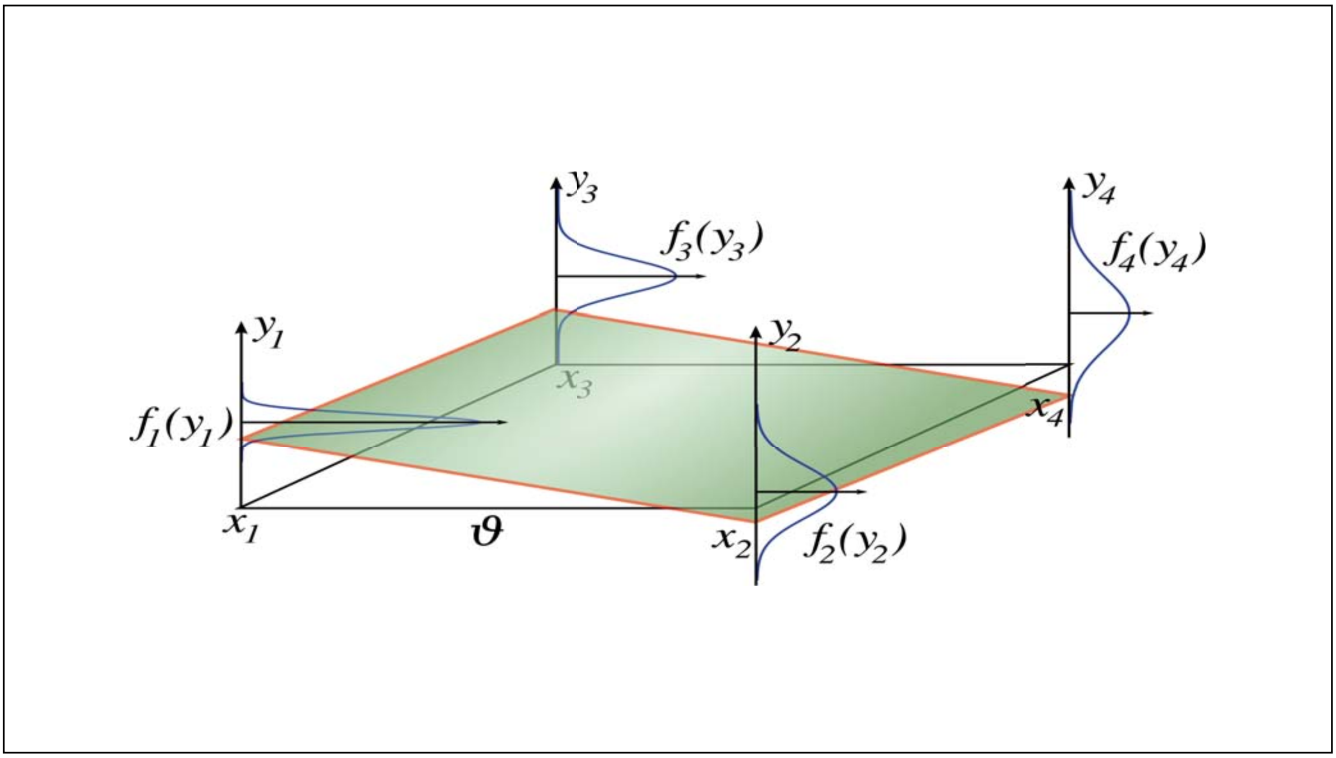
y_3
$f_3(y_3)$
y_4
$f_4(y_4)$
y_1
y_2
x_3
$f_1(y_1)$
x_4
x_1
ϑ
x_2
$f_2(y_2)$



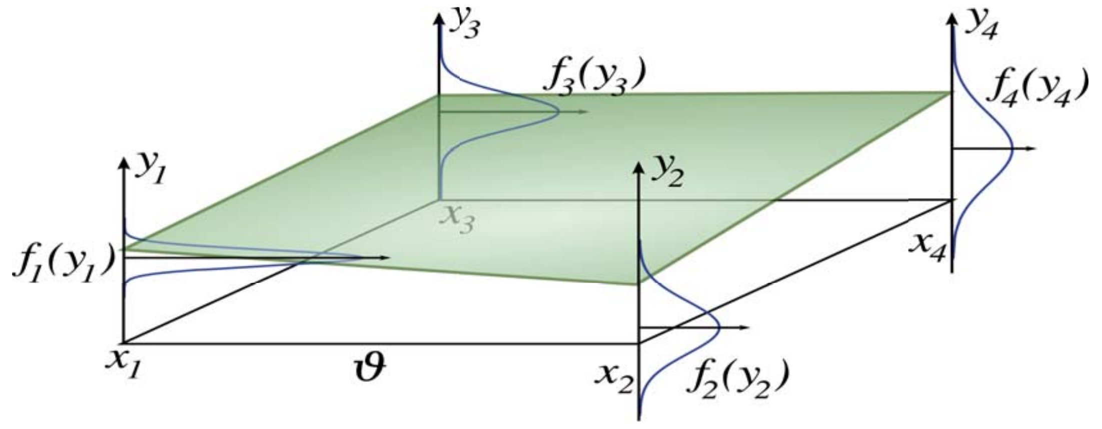
y_3
$f_3(y_3)$
y_4
$f_4(y_4)$
y_1
y_2
x_3
$f_1(y_1)$
x_4
x_1
ϑ
x_2
$f_2(y_2)$

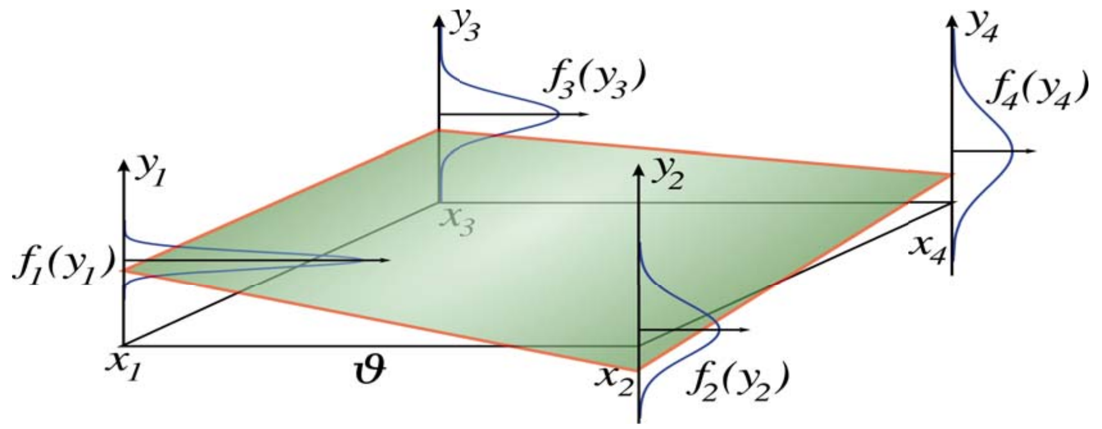


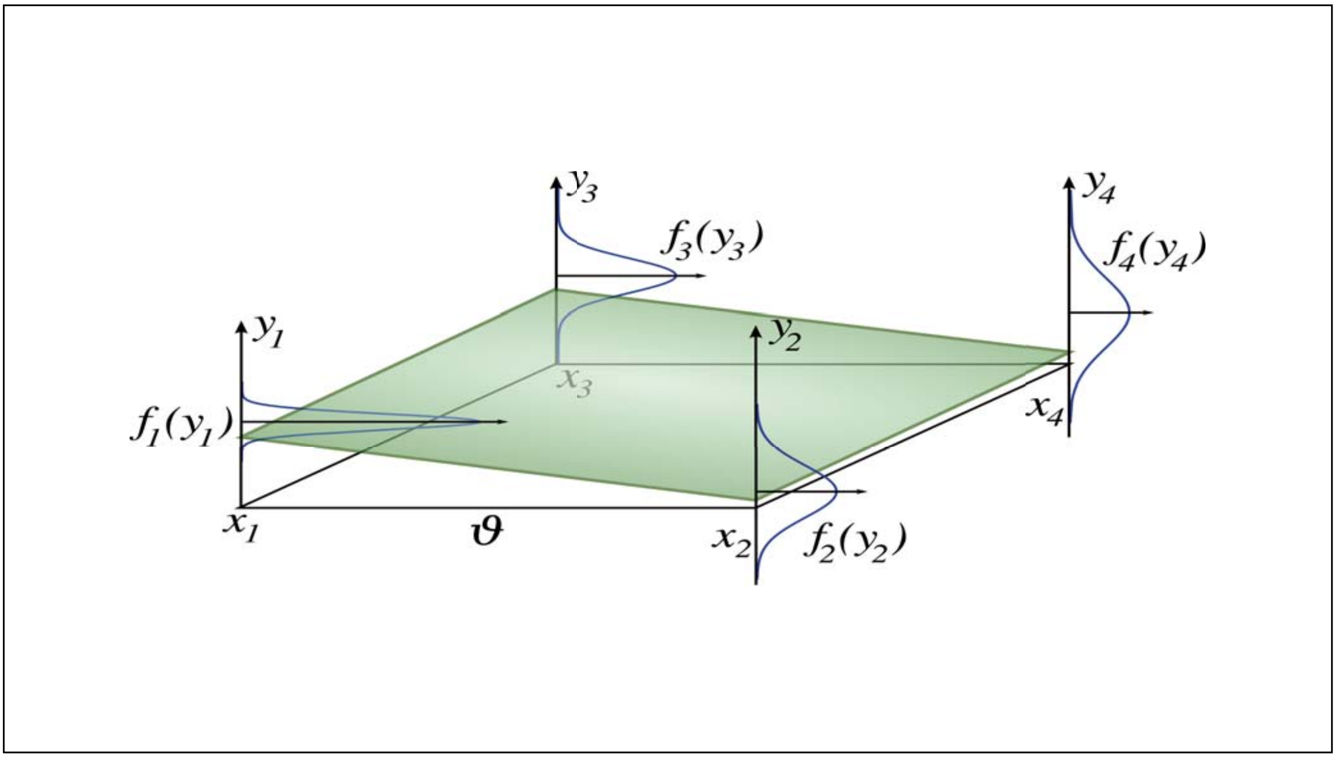
y_3
$f_3(y_3)$
y_4
$f_4(y_4)$
y_1
y_2
x_3
$f_1(y_1)$
x_4
x_1
ϑ
x_2
$f_2(y_2)$

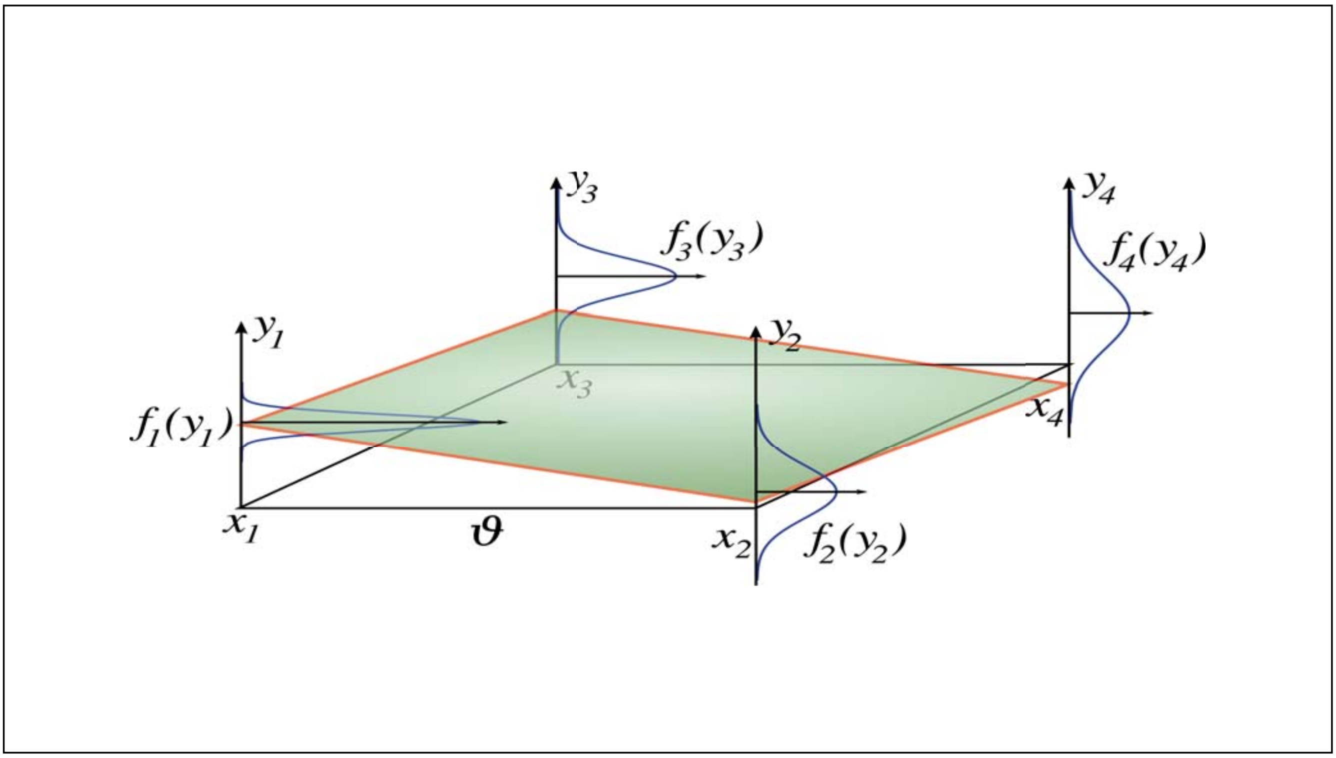


y_3
$f_3(y_3)$
y_4
$f_4(y_4)$
y_1
y_2
x_3
$f_1(y_1)$
x_4
x_1
ϑ
x_2
$f_2(y_2)$


y_3
$f_3(y_3)$
y_4
$f_4(y_4)$
y_1
y_2
x_3
x_4
$f_1(y_1)$
x_1
ϑ
x_2
$f_2(y_2)$


y_3
$f_3(y_3)$
y_4
$f_4(y_4)$
y_1
y_2
x_3
x_4
$f_1(y_1)$
x_1
ϑ
x_2
$f_2(y_2)$

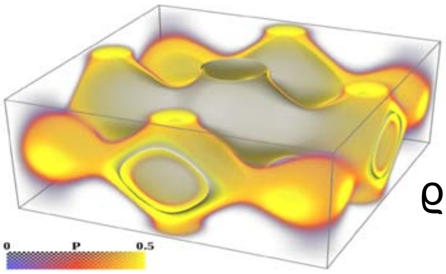

y_3
$f_3(y_3)$
y_4
$f_4(y_4)$
y_1
y_2
x_3
x_4
$f_1(y_1)$
x_1
ϑ
x_2
$f_2(y_2)$


y_3
$f_3(y_3)$
y_4
$f_4(y_4)$
y_1
y_2
x_3
x_4
$f_1(y_1)$
x_1
ϑ
x_2
$f_2(y_2)$

Algorithm & Implementation

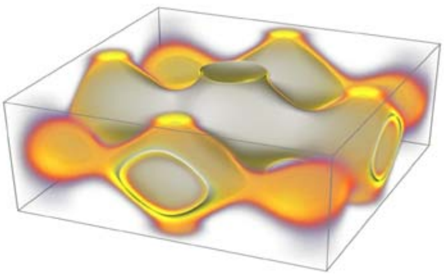
for each cell c {
    $L_c \leftarrow \text{CholeskyDecomposition}(\Sigma_c)$
    $\#\text{crossings} \leftarrow 0$
    for $1 \ldots \#\text{samples}$ {
        $\mathbf{y} \leftarrow$ random numbers $y_1 \ldots y_m \sim \mathcal{U}(0,1)$
        $\mathbf{y} \leftarrow \text{BoxMullerTransform}(\mathbf{y})$
        $\mathbf{y} \leftarrow L_c\,\mathbf{y} + \mu_c$
        $\text{if}(\text{crossing}_\vartheta(\mathbf{y}))\ \#\text{crossings} \leftarrow \#\text{crossings} + 1$
    }
    $\text{Prob}_c \leftarrow \#\text{crossings}/\#\text{samples}$
}

Impact of Spatial Correlations

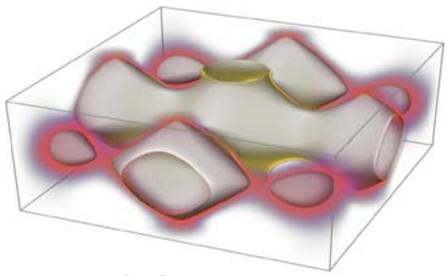


ϱ=0.00

synthetic data



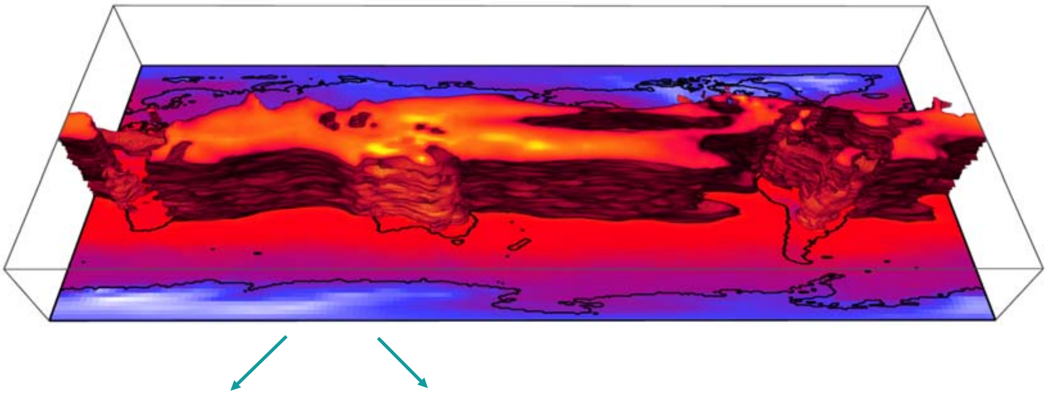
ϱ=0.65



ϱ=0.95

Climate Simulation

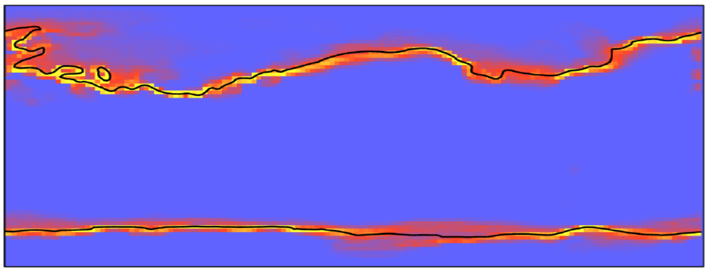
Data courtesy of ECMWF



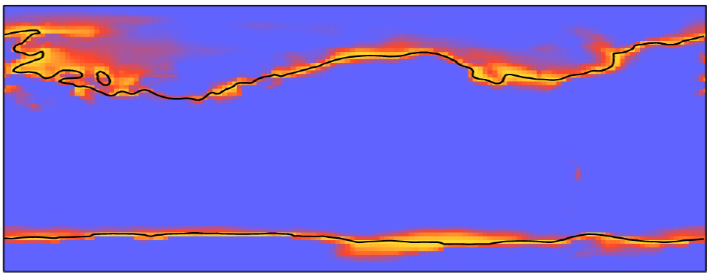
$\hat{\mu}_i$ $\widehat{\mathrm{Cov}}_{i,j}$

Isotherm of Climate Simulation

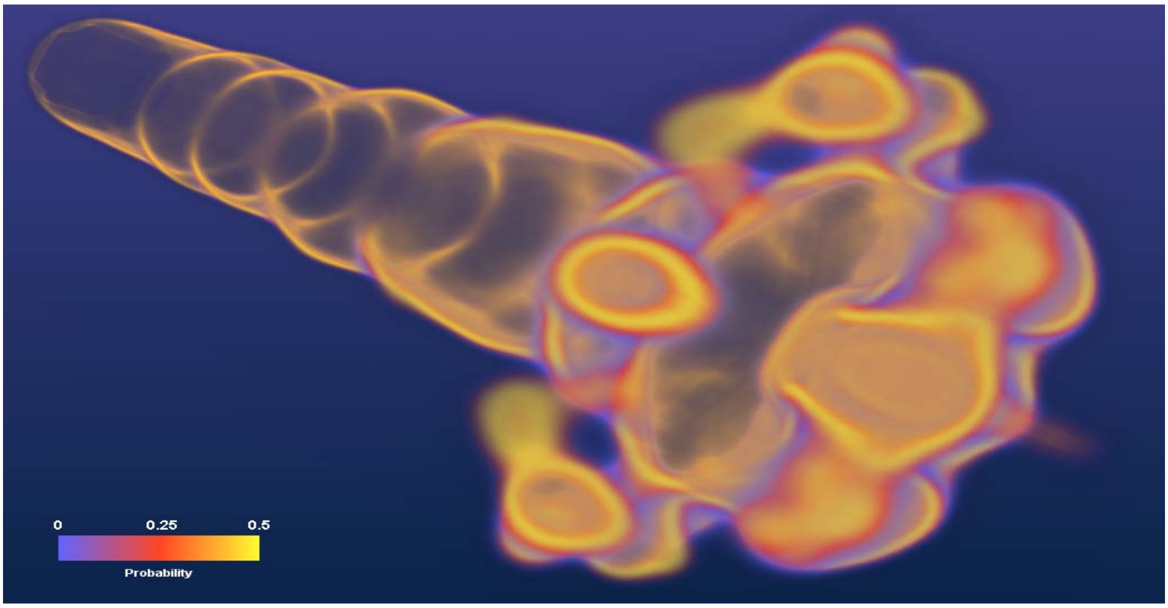
spatial correlations considered



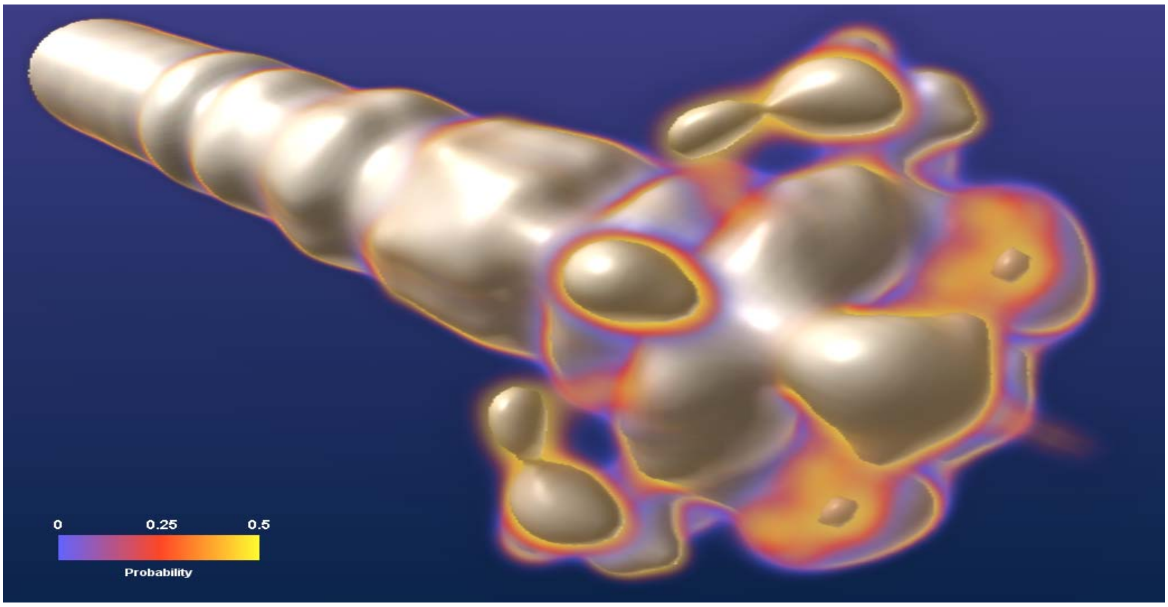
not considered



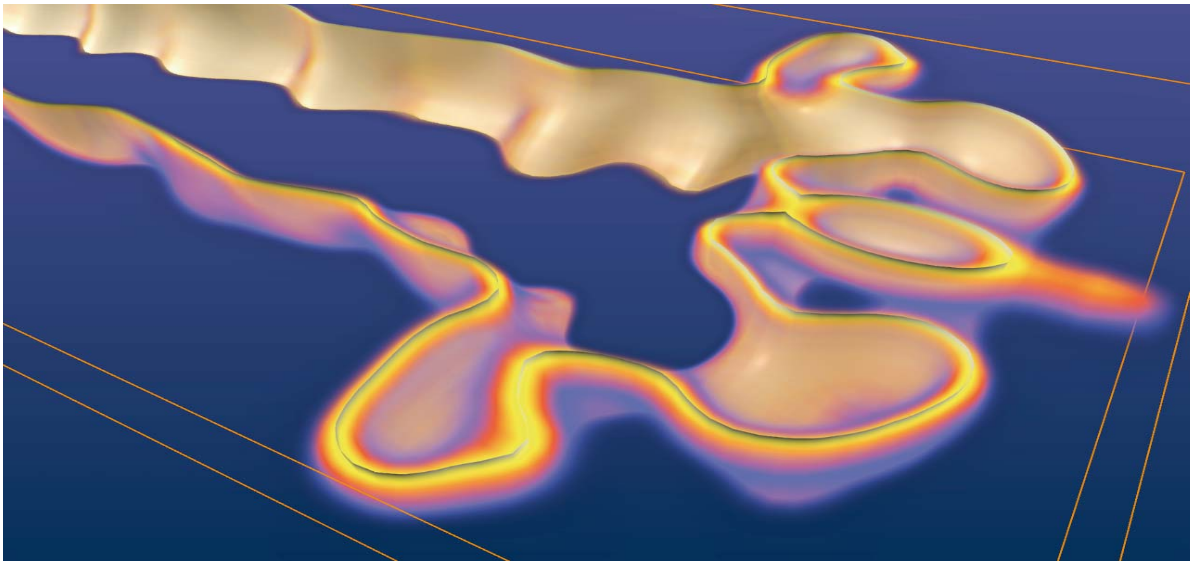
Fuel Injection Data Set + Artificial Noise: Uncertain Level Set



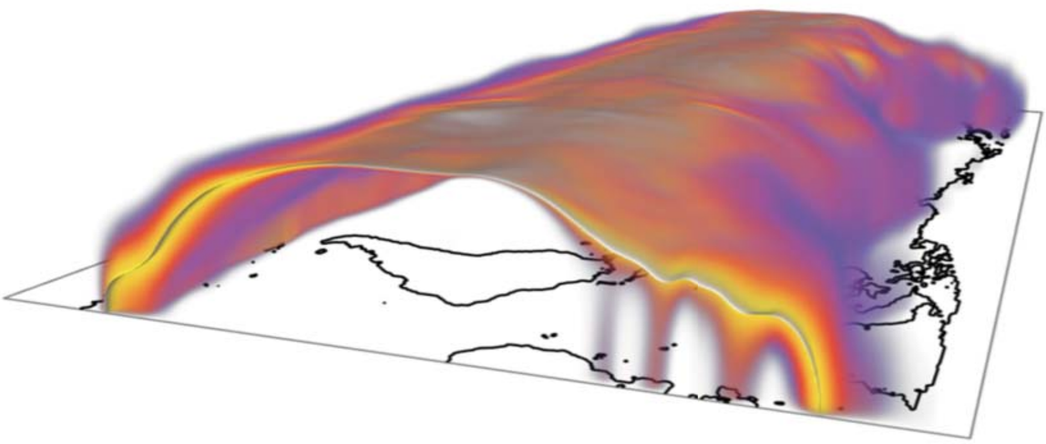
Fuel Injection Data Set + Artificial Noise: Uncertain Level Set



Fuel Injection Data Set + Artificial Noise: Uncertain Level Set

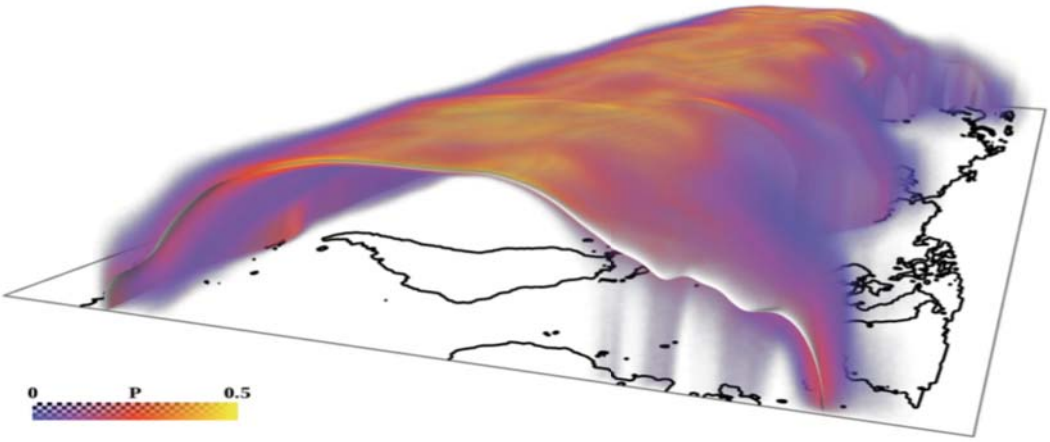


Application Example: Isotherm of Climate Simulation



spatial correlations not considered

Application Example: Isotherm of Climate Simulation



spatial correlations considered

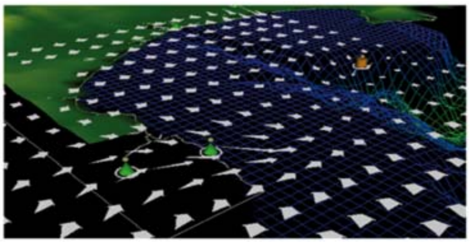
Local Features in Uncertain Vector Fields

Probabilistic local features in uncertain vector fields with spatial correlation

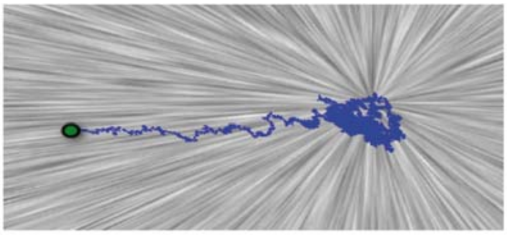
Christoph Petz, Kai Pöthkow, HCH

Computer Graphics Forum 31:3, 2012, pp. 1325-1334

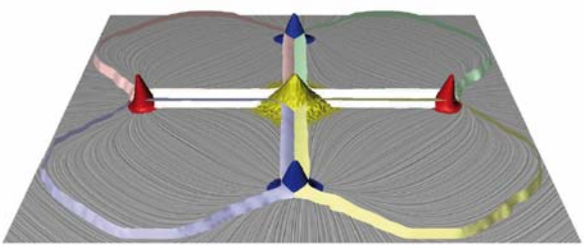
Previous Work



Wittenbrink, Pang & Lodha
Glyphs for visualizing uncertainty in vector fields
TVCG, 1996

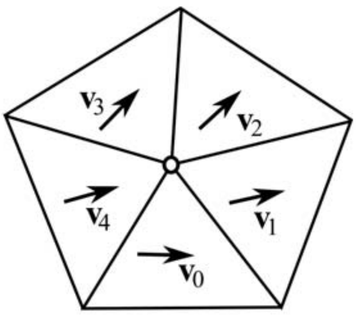


Friman, Hennemuth, Harloff, Bock, Markl & Peitgen,
Probabilistic 4D blood flow tracking and uncertainty estimation,
Medical Image Analysis, 2011

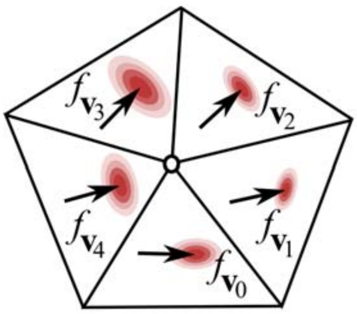


Otto, Germer, Hege & Theisel
Uncertain 2D Vector Field Topology,
Eurographics 2010

Discretized Vector Fields



crisp vector field



uncertain vector field

Tasks

- Define a model to **represent** uncertain vector fields considering spatial correlation
- Establish a framework for local **probabilistic** feature extraction from vector fields
- Estimate probabilities for the existence of **critical points** and **vortex cores**

Uncertain Vector Fields

- Again modeled as **discrete random field**
- For simplicity: Normal distributions $\mathbf{Y} \sim \mathcal{N}_n(\mu, \Sigma)$

$$n = Nd$$

# of sample points (N) — *dimension of vectors* (d)

$$\mathbf{Y}(\mathbf{y}) = \frac{1}{(2\pi)^{n/2}\det(\Sigma)^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{y}-\mu)^T\Sigma^{-1}(\mathbf{y}-\mu)\right)$$

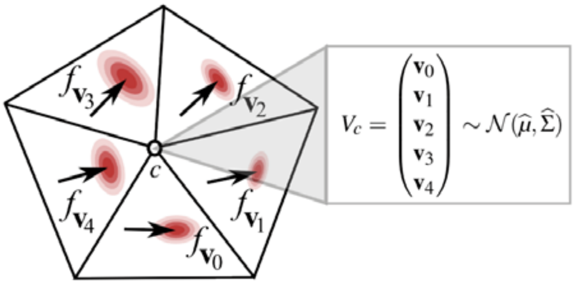
$$\mu = [\mathrm{E}(Y_1), \mathrm{E}(Y_2), \ldots, \mathrm{E}(Y_n)]$$

$$\Sigma = [\mathrm{Cov}(Y_i, Y_j)]_{i=1,2,\ldots,n;j=1,2,\ldots,n}.$$

Marginalization

Local features can be identified

- at each cell (and its neighborhood)
- using local marginal distributions



$$\widehat{\mu} : C_\eta \to \mathbb{R}^{K_c N}$$

$$\widehat{\Sigma} : C_\eta \to \mathbb{R}^{K_c N \times K_c N}$$

- Feasible for Gaussian fields only

Probabilistic Feature Extraction

Feature indicator

$$I : C_\eta \times \mathbb{R}^{K_c N} \to \{0, 1\}$$

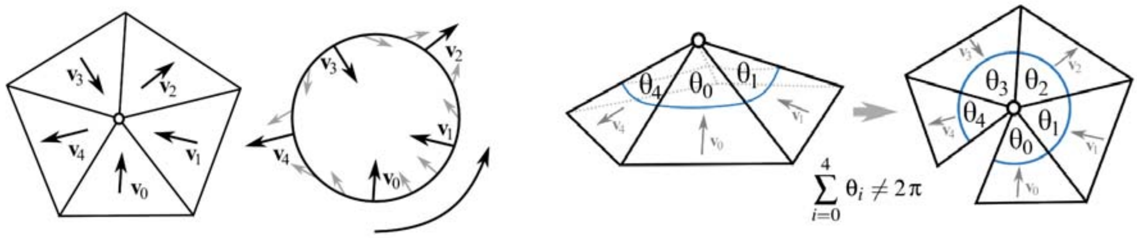
Feature probability

$$\mathrm{P}(c) = \int_D f_c(\mathbf{v})\, \mathrm{d}\mathbf{v} = \int_{\mathbb{R}^{K_c N}} f_c(\mathbf{v})\, I(c, \mathbf{v})\, \mathrm{d}\mathbf{v} = \mathrm{E}(I(c, \cdot))$$

where $D = \{\mathbf{v} \in \mathbb{R}^{K_c N} \mid I(c, \mathbf{v}) = 1\}$

Critical Points in 2D

Compute Poincaré-index (winding number)



where
$$\mathrm{idx}(c,\mathbf{v}) = \frac{\sum_{i=0}^{K-1} \measuredangle(\mathbf{v}_i, \mathbf{v}_{(i+1)\%K})}{\sum_{i=0}^{K-1} \theta_i}$$

Critical Point Classification

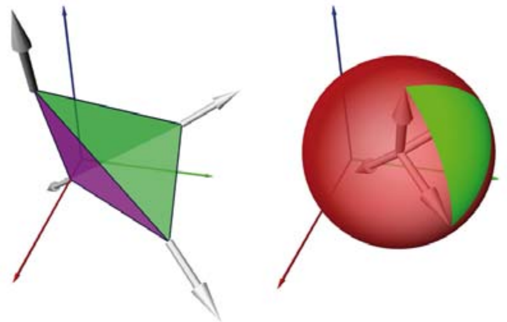
$$I_{\mathrm{source}}(c,\mathbf{v}) = \begin{cases} 1 & \mathrm{idx}(c,\mathbf{v}) > 0 \wedge \mathrm{div}(c,\mathbf{v}) > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$I_{\mathrm{sink}}(c,\mathbf{v}) = \begin{cases} 1 & \mathrm{idx}(c,\mathbf{v}) > 0 \wedge \mathrm{div}(c,\mathbf{v}) < 0 \\ 0 & \text{otherwise} \end{cases}$$

$$I_{\mathrm{saddle}}(c,\mathbf{v}) = \begin{cases} 1 & \mathrm{idx}(c,\mathbf{v}) < 0 \\ 0 & \text{otherwise} \end{cases}$$

Critical Points in 3D

- For linear tetrahedral elements: 12-dimensional random vectors have to be considered
- Compute the Poincaré-index using solid angles

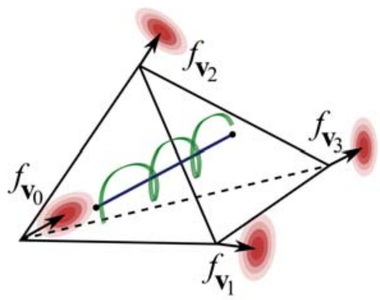


$$I_+(c,\mathbf{v}) = \begin{cases} 1 & \mathrm{idx}(c,\mathbf{v}) > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$I_-(c,\mathbf{v}) = \begin{cases} 1 & \mathrm{idx}(c,\mathbf{v}) < 0 \\ 0 & \text{otherwise} \end{cases}$$

Vortex Cores

- Indicator for vortices (Sujudi-Haimes criterion)
  - Jacobian J has 2 complex eigenvalues
  - Real eigenvector is parallel to the vector field



- J is piecewise constant → vortex cores are locally straight lines
- Compute probability for the existence of a vortex core

1st Computational Step: Empirical Parameter Estimation

- Arithmetic mean

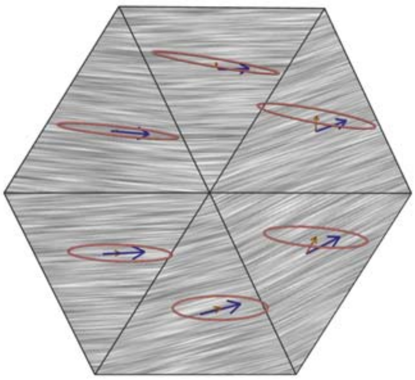
$$\hat{\mu} = \frac{1}{L} \sum_{i=1}^{L} \tilde{\mathbf{v}}_i$$

- Empirical covariance matrix

$$\hat{\Sigma} = \frac{1}{L-1} \sum_{i=1}^{L} \left(\tilde{\mathbf{v}}_i - \hat{\mu}\right) \left(\tilde{\mathbf{v}}_i - \hat{\mu}\right)^T$$

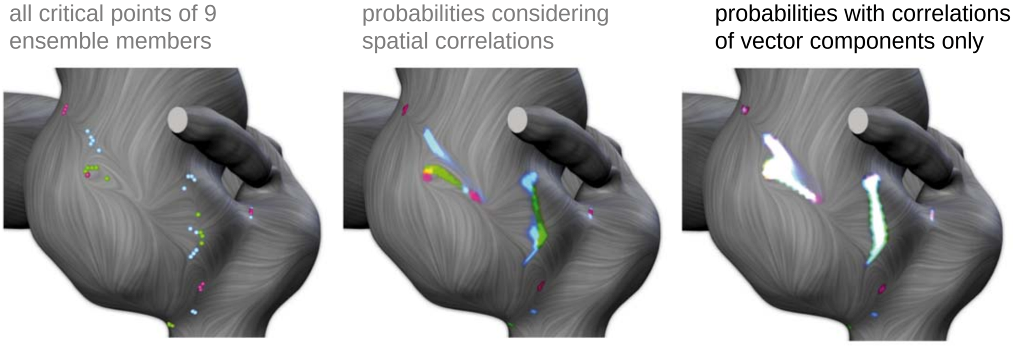
2nd Computational Step: Monte-Carlo Integration

- Compute locally correlated realizations
- Estimate feature probability using the ratio of occurrences



Critical-Point Probabilities in Wall-Shear-Stress Fields

WSS = vector field on surface



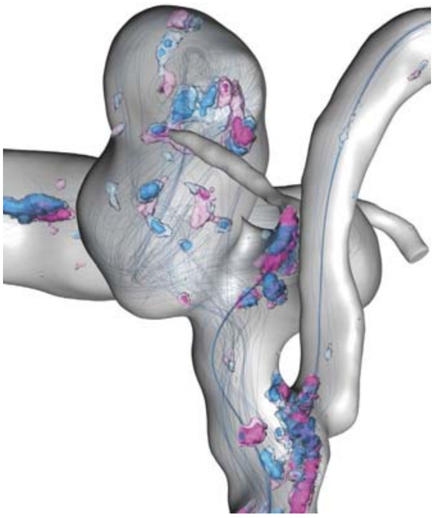
- Intensity encodes probabilities
- Color encodes type of CP: sinks in violet, sources in green
- and saddles in blue. Intensities are scaled by the probabilities.

Probabilities of CP and Swirling Motion Cores

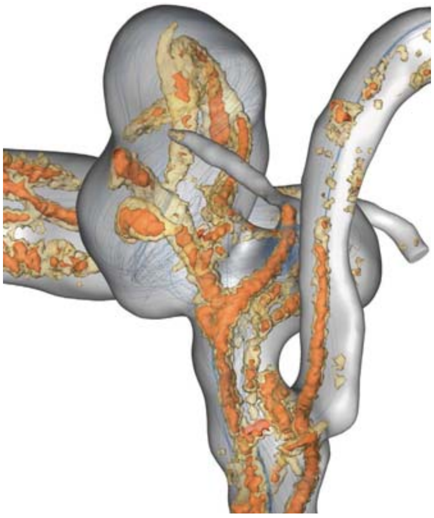
Flow features over
a full heart cycle in
a cerebral aneurysm:

Visualized by nested
semi-transparent
isosurfaces.

Streamlines of the
mean vector field
provide context.



Critical point probabilities with
Poincaré index > 0 (blue)



Probabilities for swirling
motion cores.

Research Questions in Uncertainty Vis

We Need to Understand …

Uncertainty representations

- Intervals ➔ interval computing
- Probabilities, PDFs ➔ probability theory, statistics
- Fuzzy sets ➔ soft computing
- Dempster-Shafer model ➔ evidence theory
- Possibility model ➔ possibility theory

We Need to Understand …

Reasoning under uncertainty + decision support

- Formal reasoning → statistical inference
  → uncertainty in AI
- Defuzzification, decision taking → risk & decision theory

To be Developed in Visualization

You (1986-)

Insert your photo here

- UQ in the visualization pipeline
- Fuzzy analogues of crisp features, UQ for features
- Visual mapping of uncertain / fuzzy data
- Evaluation of uncertainty representations, perceptual / cognitive efficiency
- Visual support for data processing techniques:
  data aggregation, ensemble analysis, …
- Visual support for de-fuzzification
- Visual support in decision making

Conclusion

- Uncertain iso-surfaces, critical points and vortex cores
  - reveals information not visible before
- We (still) rely on assumption of normal distribution
  - arbitrary number of realizations possible
  - more details than with limited number of realizations
- Most important research questions
  - visual mapping
  - non-Gaussian random fields
- Future of Uncertainty Vis

Thank you very much for your attention !

www.zib.de/visual

Slides: http://bit.ly/QmQfqv
http://www.cg.tuwien.ac.at/research/publications/2012/VisWeek-Tutorial-2012-Uncertainty