On the Accuracy of Homology Computations for Nodal Domains

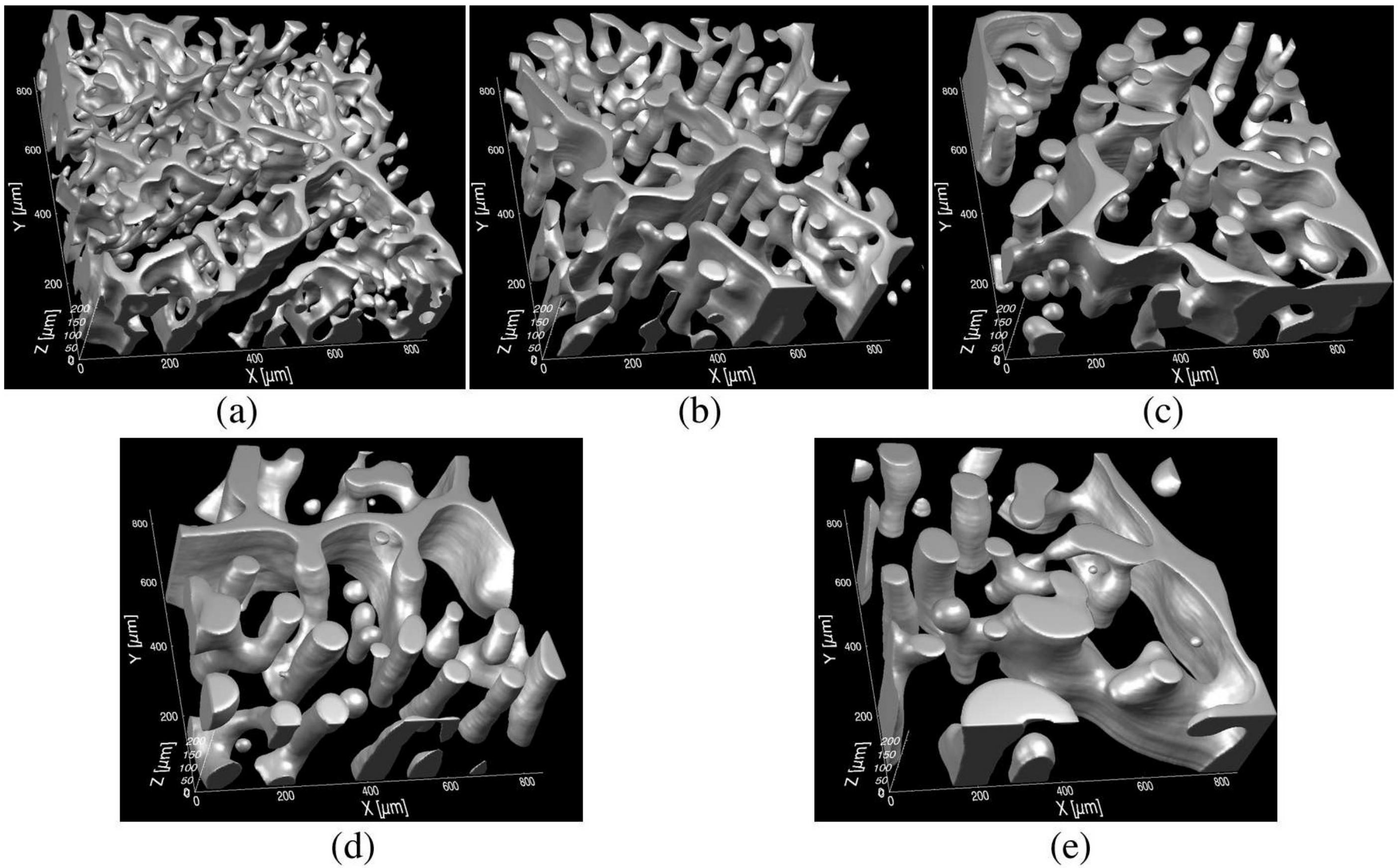
Thomas Wanner

wanner@math.gmu.edu

Department of Mathematical Sciences
George Mason University
Fairfax, Virginia 22030, USA

Phase Separation and Transient Patterns

Quenching of homogeneous binary or multi-component alloys may lead to phase separation generating complicated microstructures. The resulting patterns are generally a transient phenomenon and evolve with time.



(Figure from Mendoza, Alkemper, Voorhees (2003))

Models of Cahn-Hilliard Type

A variety of phenomenological models for such processes have been proposed over the years:

- The classical model is due to Cahn & Hilliard (1958):

$$u_t = -\Delta(\varepsilon^2 \Delta u + f(u))$$

- Cook (1970), Langer (1971): Inclusion of stochastic effects leads to the Cahn-Hilliard-Cook model:

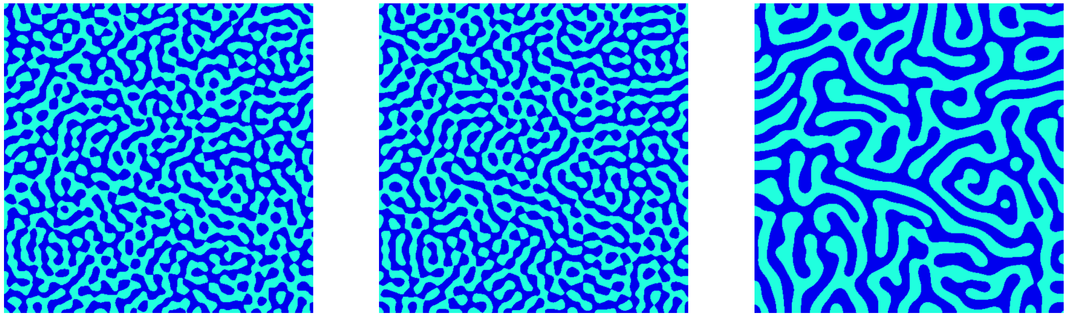
$$u_t = -\Delta(\varepsilon^2 \Delta u + f(u)) + \sigma_\varepsilon \cdot \xi$$

- Novick-Cohen (1988): Inclusion of frictional inter-phase forces leads to the viscous Cahn-Hilliard model:

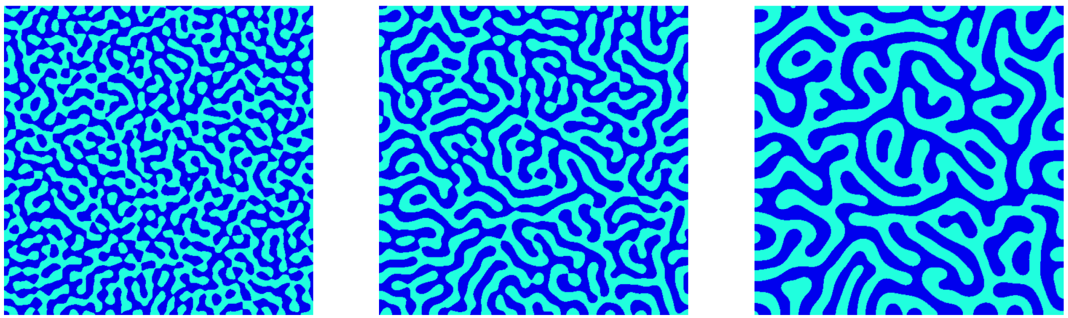
$$\beta \cdot u_t - (1-\beta) \cdot \varepsilon^2 \Delta u_t = -\Delta(\varepsilon^2 \Delta u + f(u))$$

Cahn-Hilliard-Cook Nodal Domains

Cahn-Hilliard Model with $\varepsilon = 0.005$ and total mass 0:



Cahn-Hilliard-Cook Model with $\varepsilon = 0.005$, $\sigma = 0.01$ and mass 0:



The snapshots are taken at $t = 0.0004$, $t = 0.0012$, and $t = 0.0036$.

The dark regions are the nodal domains $\{u \geq 0\}$, their light complements represent the nodal domains $\{u \leq 0\}$.

Homological Analysis of Microstructures

Fundamental problems:

- How realistic are these phenomenological models?
- Do they reproduce the microstructures accurately?
- Is a meaningful quantitative assessment possible?

Algebraic topology provides quantitative information on complex objects:

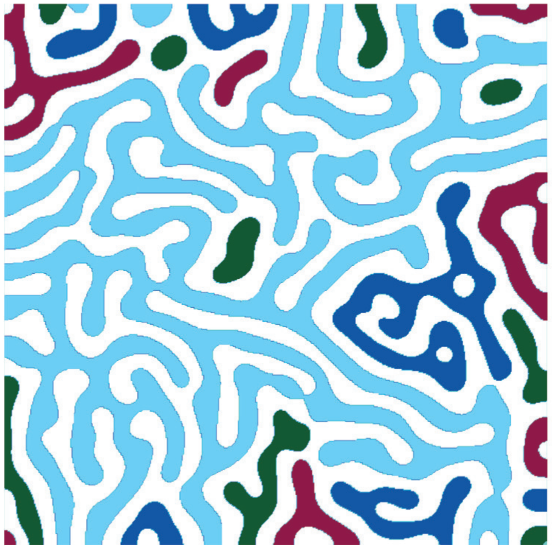
- The information is invariant under transformations which do not require cutting or gluing of the object.
- Homology groups measure the complexity of the object in any dimension.
- Betti numbers, torsion coefficients, and the Euler characteristic are coarser measures of this information.

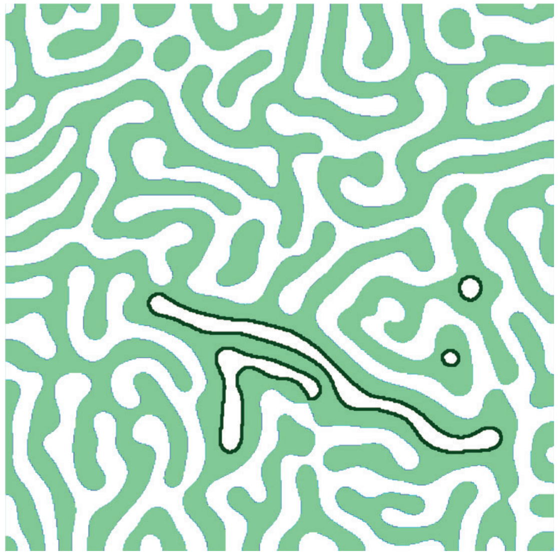
Homological Analysis of Microstructures

Gameiro, Mischaikow, W. (2005):

For total mass μ, consider the Betti numbers β_0 and β_1 of the sets

$N^+(t) = \{x \in \Omega \mid u(t,x) \geq \mu\}$ and $N^-(t) = \{x \in \Omega \mid u(t,x) \leq \mu\}$



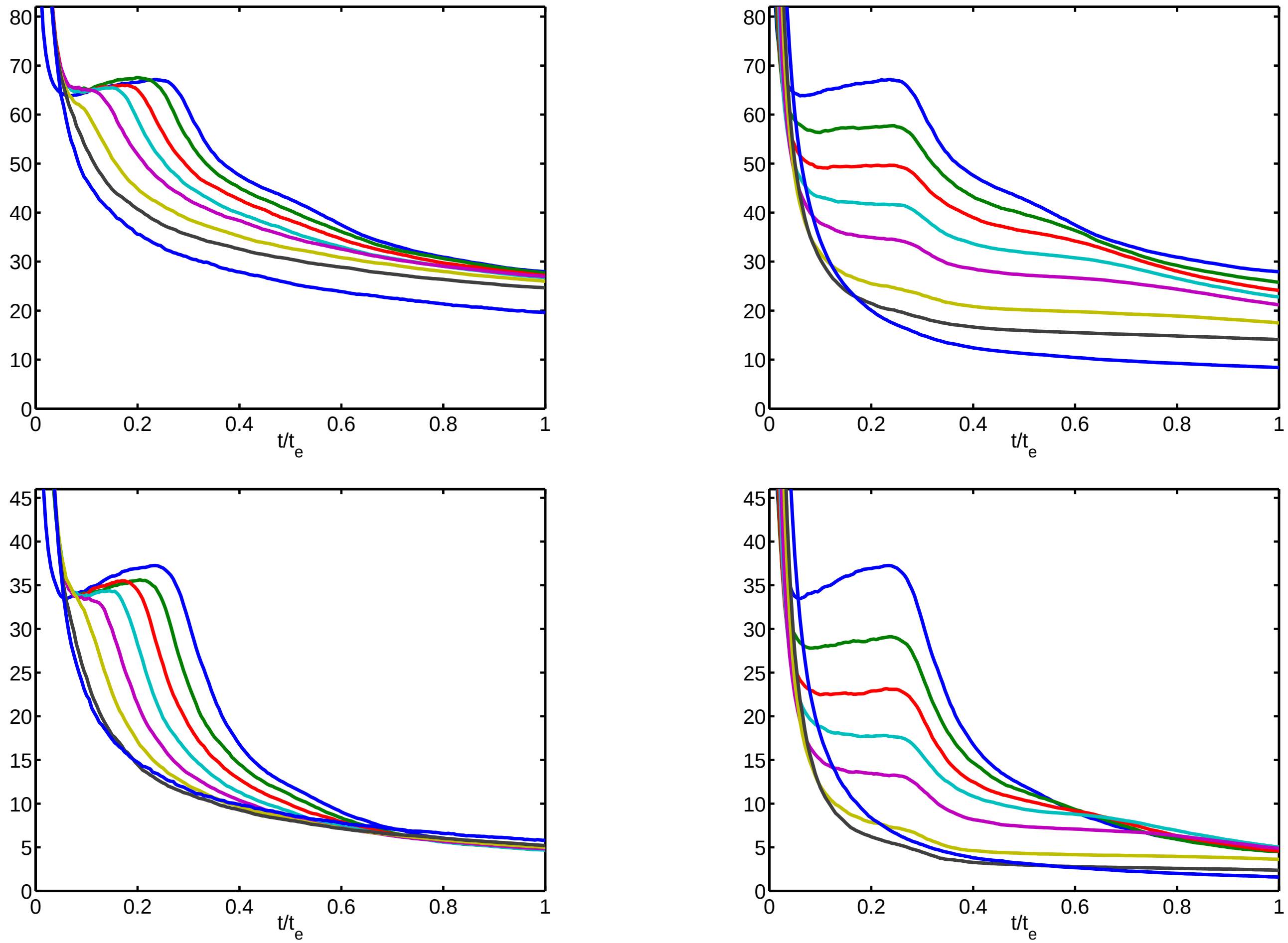


Sample set $N^+(t)$ for $\mu = 0$, $\sigma = 0$, and $t = 0.0036$.

The set has $\beta_0 = 26$ components and $\beta_1 = 4$ loops.

Averaged Betti Number Evolution

From 500 simulations for $\mu = 0$ and various values of β and σ



β_0

β_1

$N^+(t)$ for (CHC)

$N^+(t)$ for (VCH)

Homology via Discretization

Fundamental questions:

- From a mathematical point of view, the objects of interest — microstructures or patterns — are manifolds, which are often defined through level sets of differentiable functions.
- To make these objects amenable to a computational treatment, it is necessary to introduce some sort of finite discretization.
- Yet, how can one be sure that the computational results yield the correct homology of the underlying geometric object?

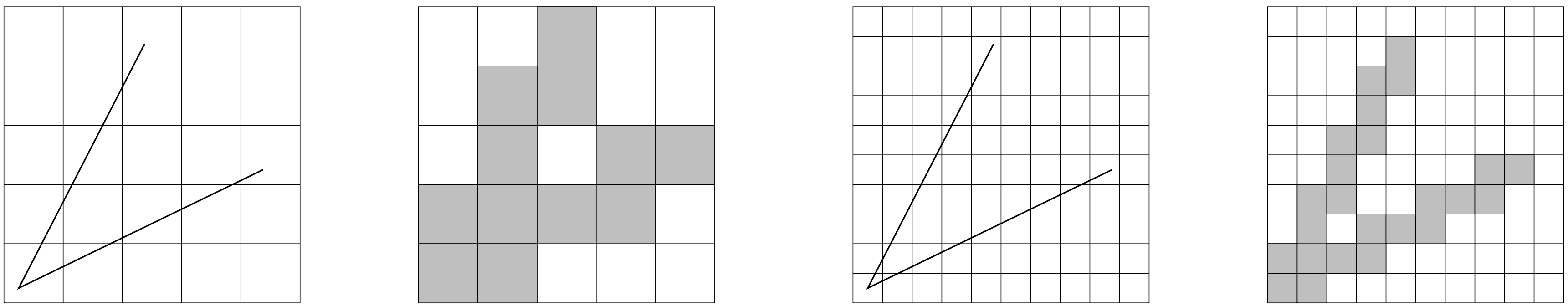
Is it enough to choose a sufficiently fine discretization?
If so, can we determine the correct discretization size a-priori?

Errors Caused by Discretization Effects

Problem:

Certain errors in homology computations which are caused by discretization effects persist even for finer discretizations — and are therefore more or less unavoidable.

Example from *Computational Homology* by Kaczynski, Mischaikow, and Mrozek (2003):



(a) (b) (a) (b)

Probabilistic Approach for Manifolds

Is it possible to determine the likelihood of success or failure of performing a homology computation with a given discretization?

Niyogi, Smale, Weinberger (2004): Homology of manifolds

- Choose sample points $x_1, \ldots, x_n$ from the given manifold $\mathcal{M}$ in $\mathbb{R}^d$ according to the uniform probability measure on it.
- For some $\varepsilon > 0$, consider the union of all balls with radius ε and centers at the points x_k, i.e.,

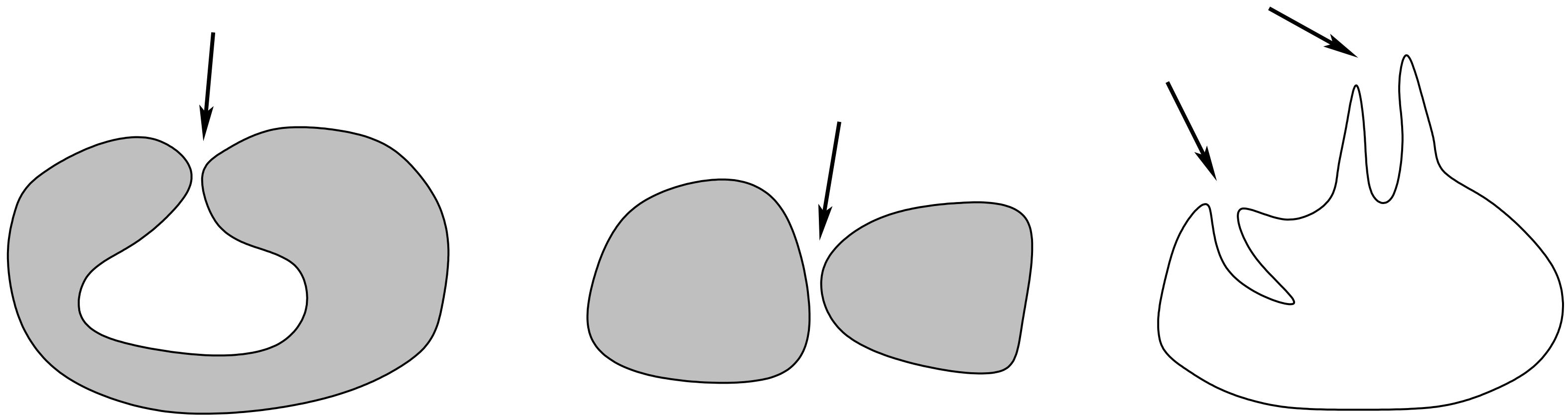
$$\mathcal{U} = \bigcup_{k=1}^{n} B_\varepsilon(x_k)$$

- Using the nerve lemma one can show that for suitable x_k and suitable ε the homologies of $\mathcal{U}$ and $\mathcal{M}$ coincide.

Probabilistic Approach for Manifolds

Crucial manifold parameter: Condition number $1/\tau$

- The inverse condition number τ is the largest number such that the open normal bundle about $\mathcal{M} \subset \mathbb{R}^d$ of radius r is embedded in $\mathbb{R}^d$ for all $r < \tau$.
- The condition number encodes both local curvature information and global separation information.



(a) (b) (c)

Probabilistic Approach for Manifolds

Main result in Niyogi, Smale, Weinberger (2004):

Let $\mathcal{M}$ be a compact manifold in $\mathbb{R}^d$ with condition number $1/\tau$, and let $x_1, \ldots, x_n \in \mathcal{M}$ be drawn in i.i.d. fashion according to the uniform probability measure on $\mathcal{M}$. Let $0 < \varepsilon < \tau/2$ and let $\mathcal{U}$ denote the union of the balls $B_\varepsilon(x_k)$, $k = 1, \ldots, n$. Then for all

$$\boxed{n > \beta_1 \cdot \ln \frac{\beta_2}{\delta}}$$

the homology of $\mathcal{U}$ equals the homology of $\mathcal{M}$ with probability at least $1 - \delta$. The constants are given by

$$\theta = \arcsin \frac{\varepsilon}{2\tau}\,, \quad \beta_1 = \frac{vol(\mathcal{M})}{\cos^d \theta \cdot vol(B_\varepsilon)}\,, \quad \beta_2 = \frac{vol(\mathcal{M})}{\cos^d \theta \cdot vol(B_{\varepsilon/8})}\,.$$

Probabilistic Approach for Manifolds

Implications of Niyogi, Smale, Weinberger (2004):

- The explicit probability estimate depends on the sample size and on the central manifold parameter which relates to its curvature and global separation.
- The result provides a-priori information on choosing a suitable discretization size — the number of points in the random sample, if the condition number can be estimated.
- The probabilistic aspect is introduced by choosing a random sample of points on the manifold.

Mischaikow, Nanda (2006):
Extension of the above result to cover the homology of maps between Riemannian manifolds.

Homology of Nodal Domains

Practical considerations:

- For simulations such as the ones described earlier, the function values are known only on a fixed regular grid which is determined by the numerical method.
- The nodal domains are not given directly, only implicitly.
- Determining or estimating the condition number of the nodal domains $\{u \geq 0\}$ and $\{u \leq 0\}$ of a function u seems difficult.
- The condition number vanishes whenever the topology changes.
- On the other hand, there is a natural notion of randomness intrinsic to the problem:
  - Random ensemble of initial conditions,
  - Stochastic evolution equation.

Random Fourier Series

Typical situation:

- For evolution equations (deterministic or stochastic) with random ensembles of initial conditions, the solution at some point in time is given as a random Fourier series

$$u(x,\omega) = \sum_{k=0}^{\infty} \alpha_k \cdot g_k(\omega) \cdot \varphi_k(x)$$

- The numbers α_k are real constants, the random variables g_k are independent, and the functions $\varphi_k : I \to \mathbb{R}$, $k \in \mathbb{N}_0$, form a complete orthogonal set in the considered function space.
- We are interested in the homology of the nodal domains

$$N^{\pm}(\omega) = \{x \in I \ : \ \pm u(x,\omega) \geq 0\}$$

Homology via Discretization in 1D

Computing the homology of nodal domains:

- Let $I = [a, b] \subset \mathbb{R}$ and consider the random nodal domains

$$N^{\pm}(\omega) = \{x \in I = [a, b] \; : \; \pm u(x, \omega) \geq 0\}$$

- Consider the discretization of I of size M given by

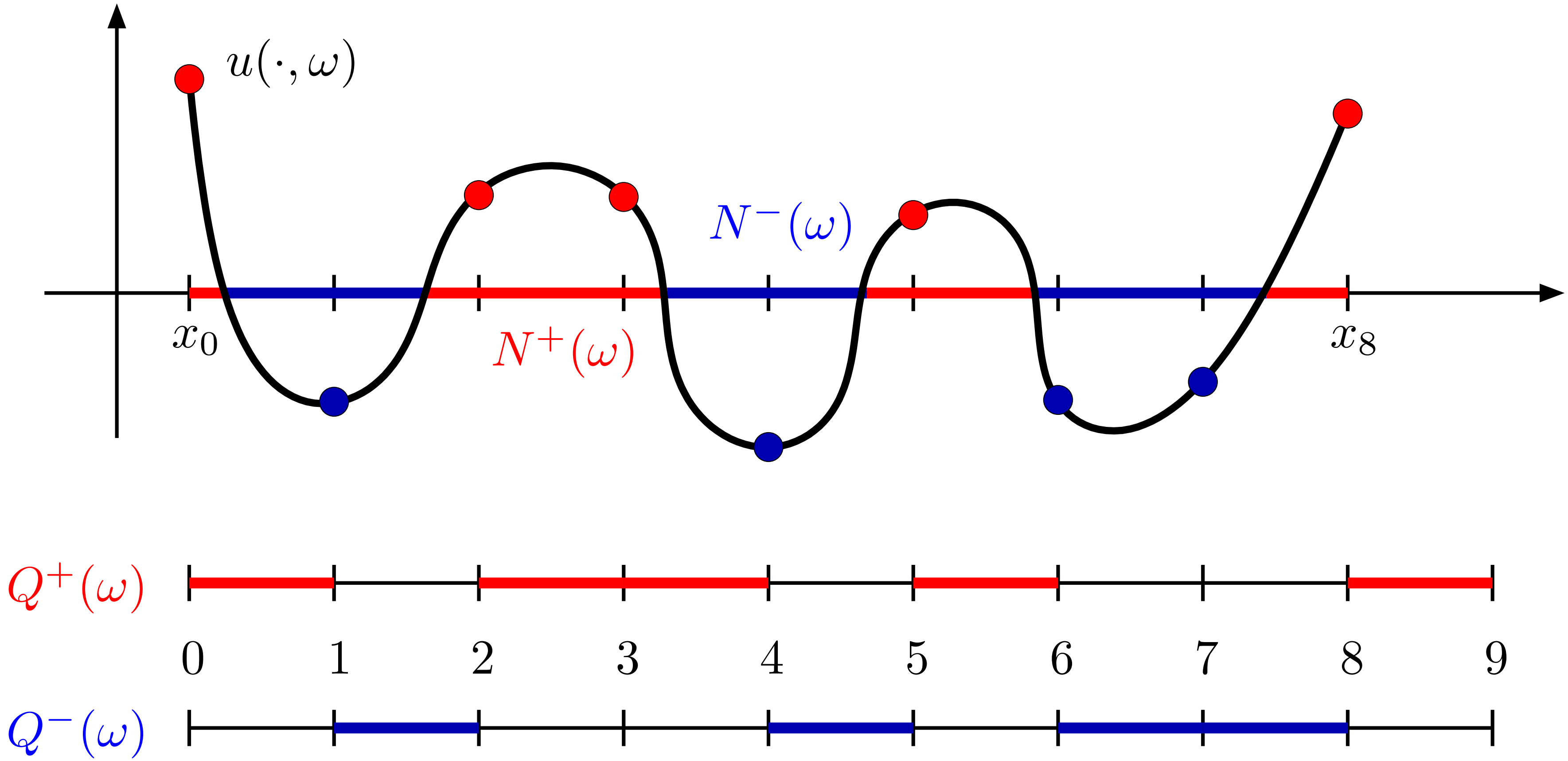
$$x_k = a + k \cdot \frac{b - a}{M} \, , \quad k = 0, \ldots, M$$

- With this discretization we associate the random cubical complexes

$$Q^{\pm}(\omega) = \bigcup \{[k, k+1] \; : \; \pm u(x_k, \omega) > 0\}$$

Homology via Discretization in 1D

$$\mathbb{P}\left\{\omega \;:\; H_*\left(N^{\pm}(\omega)\right) = H_*\left(Q^{\pm}(\omega)\right)\right\} = \;?$$



Assumptions on the Random Field

We assume that $u : I \times \Omega \to \mathbb{R}$ is almost surely continuous and

(1) For every $x \in I$ we have

$$\mathbb{P}\{u(x) = 0\} \;:=\; \mathbb{P}\{\omega \,:\, u(x,\omega) = 0\} \;=\; 0$$

(2) We have

$$\mathbb{P}\{u(\cdot) : I \to \mathbb{R} \text{ has a double zero in } I\} \;=\; 0$$

(3) For $x \in I$ and $\delta > 0$ with $x + \delta \in I$ let

$$p_\sigma(x,\delta) = \mathbb{P}\left\{\sigma \cdot u(x) \geq 0\,,\; \sigma \cdot u\left(x + \frac{\delta}{2}\right) \leq 0\,,\; \sigma \cdot u(x+\delta) \geq 0\right\}$$

Then there exists a constant $\mathcal{C}_0 > 0$ such that

$$p_{\pm 1}(x,\delta) \quad \leq \quad \mathcal{C}_0 \cdot \delta^3 \quad \text{for all} \quad x, x+\delta \in I$$

Abstract Probability Estimate

Mischaikow, W. (2006):

Consider an almost surely continuous random field $u : I \times \Omega \to \mathbb{R}$ *on the interval* $I = [a, b]$*, satisfying Assumptions (1), (2), and (3). Then the probability* P *that the homology of* $N^{\pm}(\omega)$ *is computed correctly with the discretization of size* M *satisfies*

$$P \;\geq\; 1 - \frac{8\mathcal{C}_0 (b-a)^3}{3M^2}$$

where $\mathcal{C}_0$ *denotes the constant from Assumption (3).*

For most concrete applications, Assumptions (1) and (2) can be verified easily. Only Assumption (3) usually requires some work.

Application to Periodic Random Fields

The study of evolution equations with periodic boundary conditions leads to classical random Fourier series of the form

$$u(x,\omega) \;=\; \sum_{k=0}^{\infty} a_k \cdot \left(g_{2k}(\omega) \cdot \cos(kx) + g_{2k-1}(\omega) \cdot \sin(kx)\right)$$

In particular, if we concentrate on linear evolution equations with Gaussian ensembles, then we can assume that the random variables g_k are independent and normally distributed with mean 0 and variance 1.

In this case, the series $u(x,\omega)$ is a homogeneous Gaussian random field with mean 0 and spatial covariance function

$$R(x,y) = r(x-y) = \sum_{k=0}^{\infty} a_k^2 \cdot \cos(k(x-y))$$

Result for Periodic Random Fields in 1D

Mischaikow, W. (2006):

Consider the random Fourier series u as before and assume that

$$\sum_{k=0}^{\infty} k^6 a_k^2 < \infty \quad \text{and} \quad a_{k_1} \neq 0 \,, \quad a_{k_2} \neq 0 \quad \text{for} \quad k_1 < k_2$$

Then the probability P that the homology of $N^{\pm}(\omega)$ is computed correctly with the discretization of size M satisfies

$$P \;\geq\; 1 - \frac{\pi^2}{6M^2} \cdot \frac{A_2 A_0 - A_1^2}{A_0^{3/2} A_1^{1/2}} + O\left(\frac{1}{M^4}\right)$$

where

$$A_\ell \;=\; \sum_{k=0}^{\infty} k^{2\ell} a_k^2 \;=\; \frac{1}{2\pi} \cdot \mathbb{E} \left\| D_x^\ell u \right\|_{L^2(0,2\pi)}^2$$

Result for Periodic Random Fields in 1D

What does the result imply?

- The result provides explicit probability estimates for the correctness of the homology computation.
- The probability estimate depends on the discretization size and on central parameters of the random field which relate to its smoothness properties and to its derivatives up to second order.
- The result provides a-priori information on choosing a suitable discretization size.

How sharp is this estimate?

Application: Finite Trigonometric Sums

Any random trigonometric polynomial of the form

$$u(x,\omega) \;=\; \sum_{k=1}^{N} a_k \cdot \left(g_{2k}(\omega)\cdot\cos(kx) + g_{2k-1}(\omega)\cdot\sin(kx)\right)$$

has at most $2N$ zeros. In this situation our result furnishes:

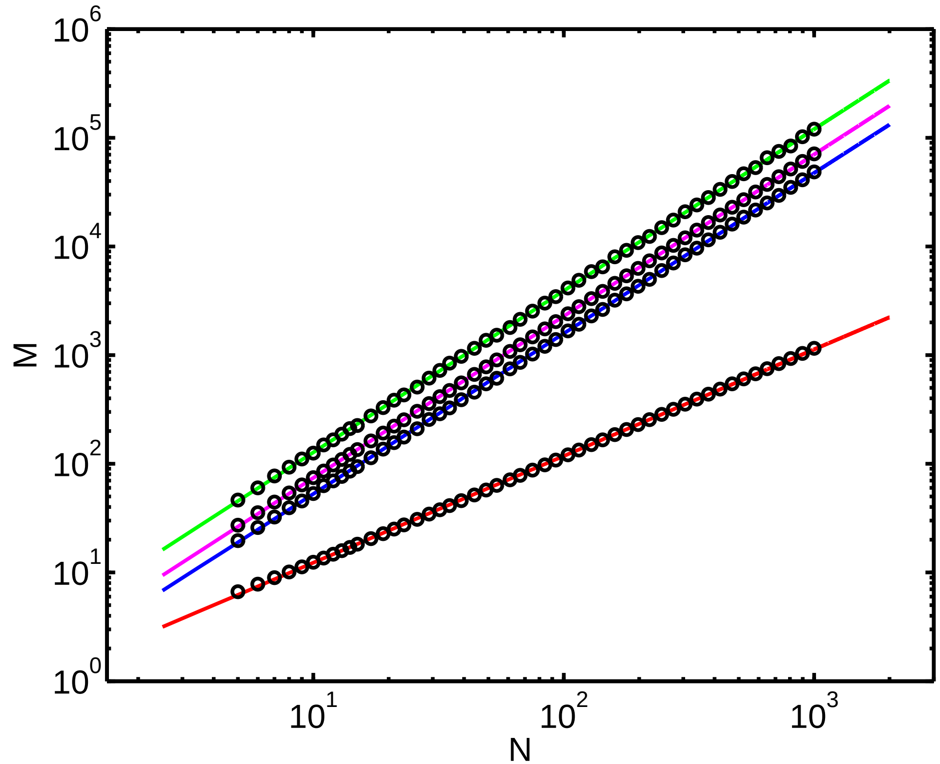
The probability P that the homology of the random nodal domains $N^{\pm}(\omega)$ is computed correctly with the discretization of size M satisfies

$$P \;\geq\; 1 - \frac{2\sqrt{3}\pi^2}{135}\cdot\frac{N^3}{M^2} + O\left(\frac{1}{M^4}\right)$$

In order to compute the homology correctly with high confidence we need to choose $M \sim N^{3/2}$.

Application: Finite Trigonometric Sums

Numerical results confirm $M \sim N^{3/2}$.



Shown are the expected number of zeros, the expected value of $2\pi/d_{\min}$, where $d_{\min}$ is the minimal distance between two zeros, the value of M for which 95% of the functions have minimal distance at least $2\pi/M$, and the value of the discretization size M for which the probability estimate yields $P = 95\%$. For each N we considered 15,000 random trigonometric sums.

Application: Linear Cahn-Hilliard Model

The solution of the linearized Cahn-Hilliard equation originating at a Gaussian random field is given by

$$u(x,\omega) \quad = \quad \sum_{k=1}^{\infty} e^{\lambda_k t} \cdot a_k \cdot \left(g_{2k}(\omega) \cdot \cos(kx) + g_{2k-1}(\omega) \cdot \sin(kx) \right)$$

where $\lambda_k = k^2(1 - \varepsilon^2 k^2)$. In this situation one obtains the probability estimate

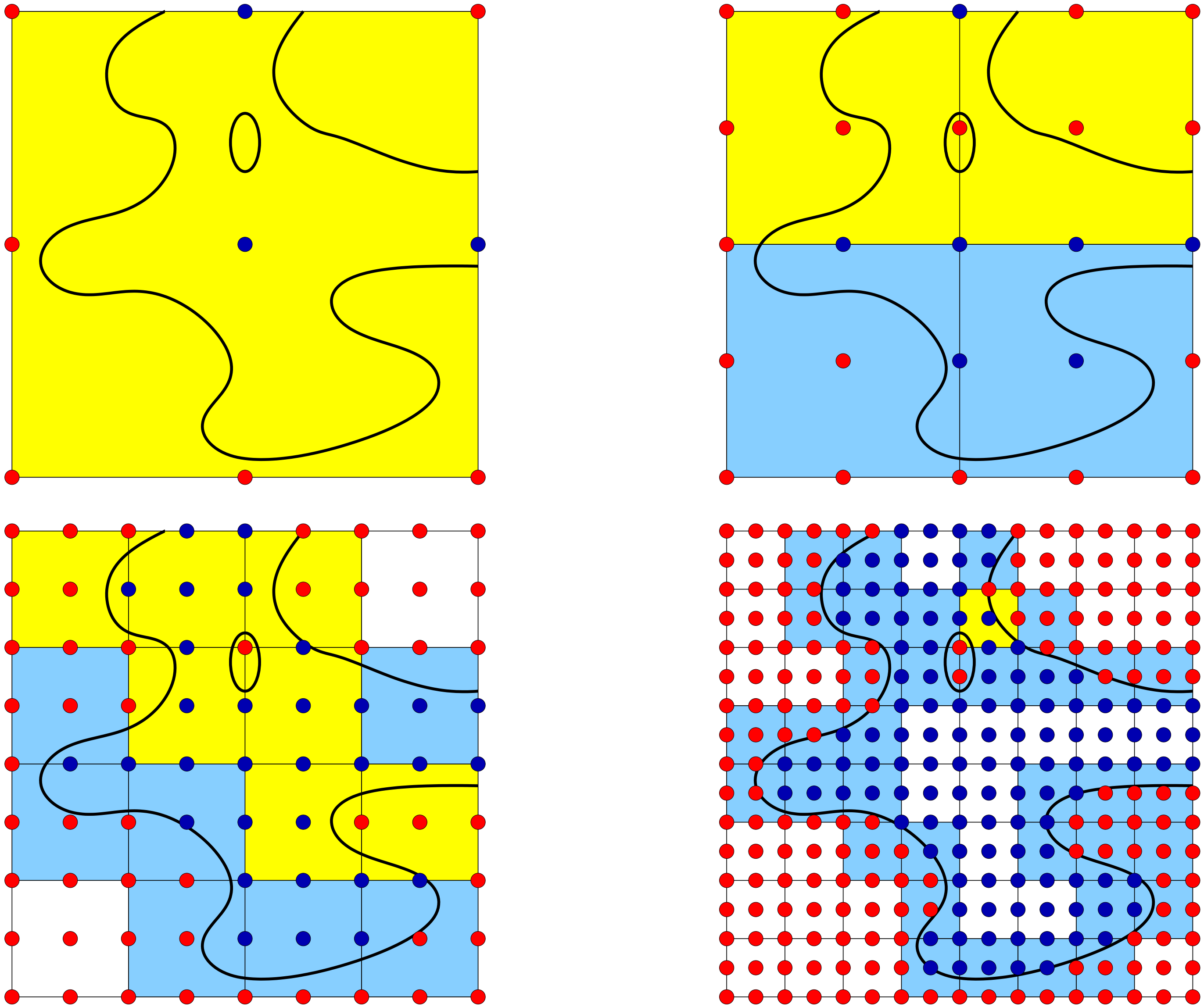
$$P \quad \geq \quad 1 - \frac{\pi^2}{6\varepsilon^3 M^2} \cdot C\left(\frac{t}{\varepsilon^2}\right) + O\left(\frac{1}{M^4}\right)$$

for some ε-independent, decreasing function C with $C(1) \approx 1/5$.

In order to compute the homology correctly with high confidence we need to choose $M \sim \varepsilon^{-3/2}$.

Towards a Two-dimensional Result

Can this result be generalized to two-dimensional domains?



Random Fourier Series in 2D

Consider a random Fourier series on $\Omega = [0, 2\pi]^2$ of the form

$$
\begin{aligned}
u(x,\omega) \quad = \quad \sum_{k,\ell=0}^{\infty} a_{k,\ell} \cdot \big(& g_{k,\ell,1}(\omega) \cos(kx_1) \cos(\ell x_2) + \\
& + g_{k,\ell,2}(\omega) \cos(kx_1) \sin(\ell x_2) \\
& + g_{k,\ell,3}(\omega) \sin(kx_1) \cos(\ell x_2) \\
& + g_{k,\ell,4}(\omega) \sin(kx_1) \sin(\ell x_2) \big)
\end{aligned}
$$

The random variables $g_{k,\ell,m}$ are independent and normally distributed with mean 0 and variance 1. There are integers $k_1, \ell_1 \in \mathbb{N}$ and $k_2, \ell_2 \in \mathbb{N}_0$ with $k_1 \neq k_2$ and $\ell_1 \neq \ell_2$ such that both $a_{k_1,\ell_1} \neq 0$ and $a_{k_2,\ell_2} \neq 0$, and in addition

$$
\boxed{\sum_{k,\ell=0}^{\infty} \left(k^6 + \ell^6\right) a_{k,\ell}^2 < \infty}
$$

Preliminary Probabilistic Result in 2D

Mischaikow, W. (2006):

The probability P that the homology of $N^{\pm}(\omega)$ is computed correctly with the discretization of size M satisfies

$$
\begin{aligned}
P \quad \geq \quad & 1 - \frac{3\pi^2}{4M} \cdot \left(\frac{A_{0,2}A_{0,0} - A_{0,1}^2}{A_{0,0}^{3/2} A_{0,1}^{1/2}} + \frac{A_{2,0}A_{0,0} - A_{1,0}^2}{A_{0,0}^{3/2} A_{1,0}^{1/2}} \right) \\
& - \quad \frac{32\pi^2}{9M^2} \cdot \frac{A_{1,1}^{3/2}}{A_{0,0}^{1/2} A_{0,1}^{1/2} A_{1,0}^{1/2}} \quad + \quad O\left(\frac{1}{M^3}\right) ,
\end{aligned}
$$

where

$$
A_{p,q} \quad = \quad \sum_{k,\ell=0}^{\infty} k^{2p} \ell^{2q} a_{k,\ell}^2 \quad = \quad \frac{1}{4\pi^2} \cdot \mathbb{E} \left\| D_{x_1}^p D_{x_2}^q u \right\|_{L^2(0,2\pi)}^2
$$

This result is suboptimal and cannot be generalized to higher dimensions! But there is room for improvement...

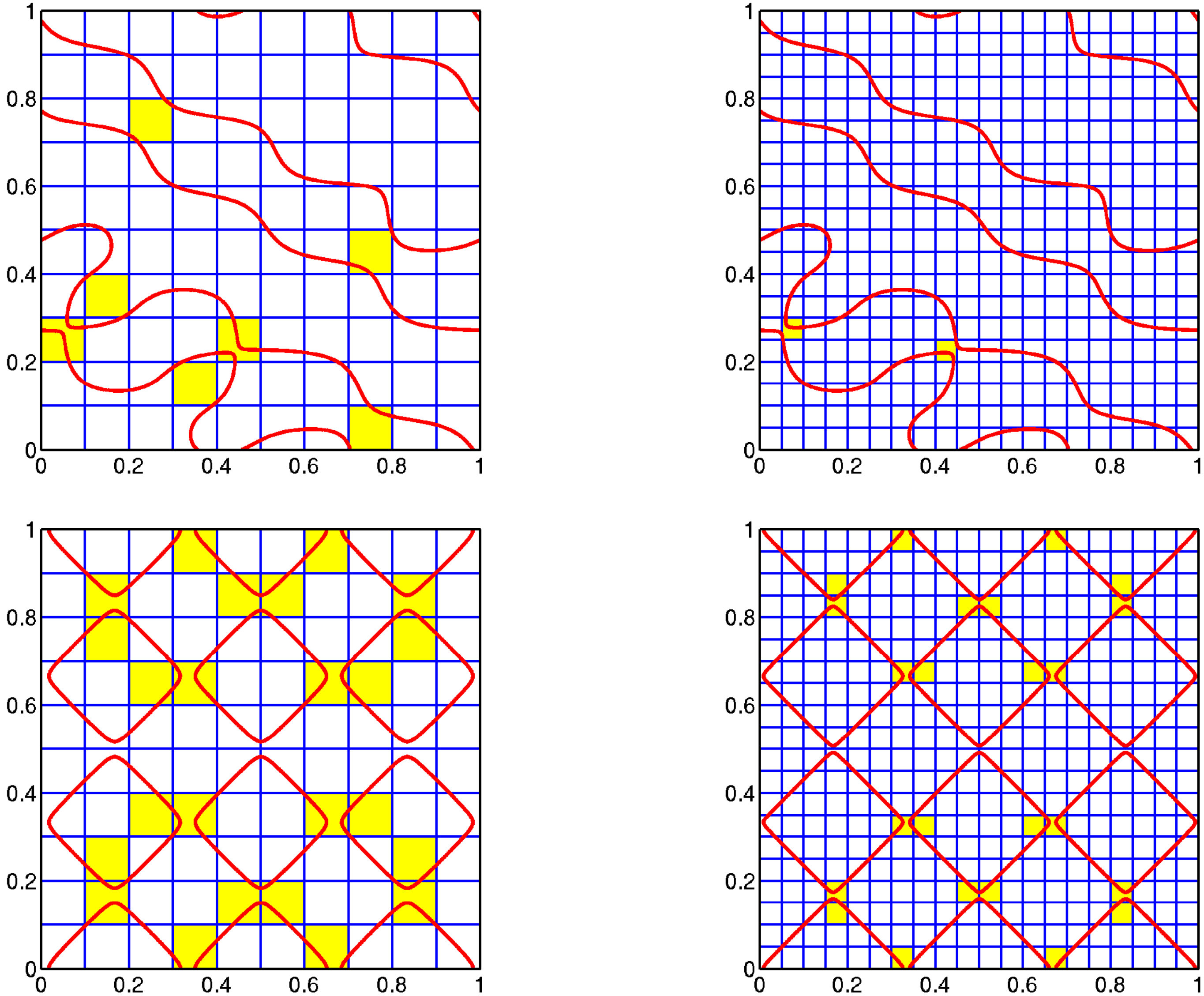
Validated Homology Computations

Day, Kalies, W. (2006):

Numerical validation of homology computations

- For nonlinear problems obtaining precise probabilistic bounds seems difficult.
- Check whether the correctness of the homology can be validated computationally.
- In some cases, validation may be impossible.
- Use interval arithmetic to obtain rigorous function value and gradient bounds.
- Preliminary results indicate that validation is possible in most cases for which the homology is correct.

Validated Homology Computations



In the yellow squares the validation was impossible.

Collaborators

- Sarah Day (Cornell University)
- Marcio Gameiro (Georgia Institute of Technology)
- Bill Kalies (Florida Atlantic University)
- Konstantin Mischaikow (Georgia Institute of Technology)

Work partially supported by DOE and NSF.