Simplicial Homology Global Optimisation

An algorithm for optimising energy surfaces

Stefan Endres

January 30, 2018

Institute of Applied Materials
Department of Chemical Engineering
University of Pretoria

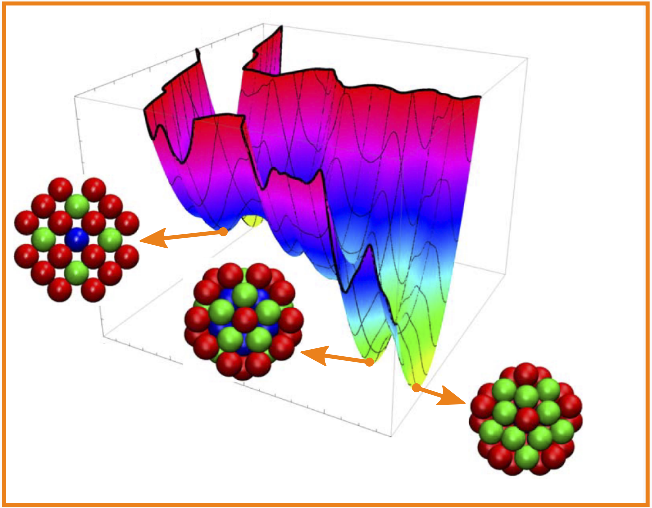
This presentation is intended for an audience of professional engineers from a diverse set of backgrounds. For researchers and experts with a strong background in optimisation theory and applied mathematics a more detailed presentation can be found at https://stefan-endres.github.io/shgo/files/shgo_slides.pdf ▶ Link

Introduction

Introduction

- Global optimisation of black-box functions
- Developed for applications on free energy (hyper-)surface problems common in chemical engineering and many other fields, examples:
  - Phase equilibria (chemical engineering)
  - Inorganic molecular structures (computational chemistry)
  - Protein folding (computational biochemistry)
  - Time independent Hamiltonian systems (quantum mechanics)
  - Equilibrium in arbitrary force models (ex. stable orbits)
- Information extracted by shgo in the limits:
  - Finds the global minimum (ex. stable equilibrium, "best" solutions)
  - Finds all other solutions (ex. corresponding to quasi-equilibrium states that have physical meaning)

Example of a free energy surface (adapted from [Wales, 2015])



Objective function statement and nomenclature

Objective function statement i

Consider a general optimisation problem of the form

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{x}), \text{ by varying } \mathbf{x} \in \mathbb{R}^n \\ \text{subject to} \quad & g_i(\mathbf{x}) \geq 0, \ \forall i = 1, \ldots, m \\ & h_j(\mathbf{x}) = 0, \ \forall j = 1, \ldots, p \end{aligned}$$

- The objective function maps an n-dimensional real space to a scalar value $f : \mathbb{R}^n \to \mathbb{R}$
- The variables $\mathbf{x}$ are assumed to be bounded
- $g_i(x)$ are the inequality constraints $\mathbf{g} : [\mathbf{l}, \mathbf{u}]^n \to \mathbb{R}^m$
- $h_j(x)$ are the equality constraints $\mathbf{h} : [\mathbf{l}, \mathbf{u}]^n \to \mathbb{R}^j$
- It is assumed that the objective function has a finite number of local minima

Objective function statement ii

for example if lower and upper bounds l_i and u_i are implemented for each variable then we have an initially defined hyperrectangle

$$\mathbf{x} \in \Omega \subseteq [\mathbf{l}, \mathbf{u}]^n \ = \ [l_1, \ u_1] \ \times \ [l_2, \ u_2] \ \times \ \ldots \ \times \ [l_n, \ u_n] \subseteq \mathbb{R}^n \qquad (1)$$

where Ω is the limited feasible subset excluding points outside the bounds and constraints.

$$\Omega = \{\mathbf{x} \in [\mathbf{l}, \mathbf{u}]^n \mid \mathbf{g}_i(\mathbf{x}) \geq 0, \forall i = 1, \ldots, m\} \qquad (2)$$

When the constraints in **g** are linear the set Ω is always a compact space.

A brief one-dimensional motivation

A brief one-dimensional motivation i

Derivative free optimisation:

- f and g are expensive black-box functions
- No derivative information available or difficult to compute
- Common strategies in global optimisation hit the maps f and g with sampling points and use the resulting geometric information of the surfaces
- Many popular approaches are based on some kind of statistical or geometric reasoning or even more simply a multi-start routine that simply passes any promising sampling points to a local minimization routine

A brief one-dimensional motivation ii

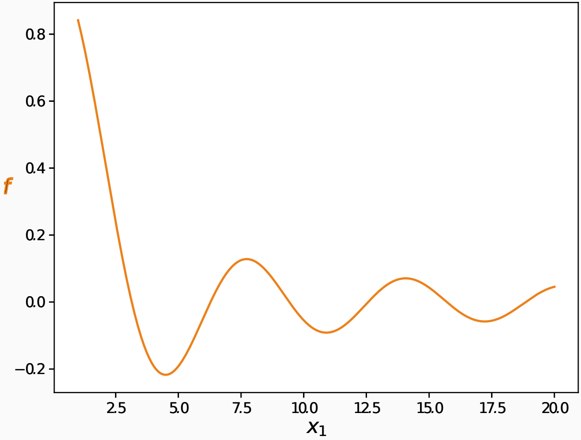


Figure 1: A 1-dimensional objective function surface $f : \mathbb{R}^1 \to \mathbb{R}$

A brief one-dimensional motivation iii

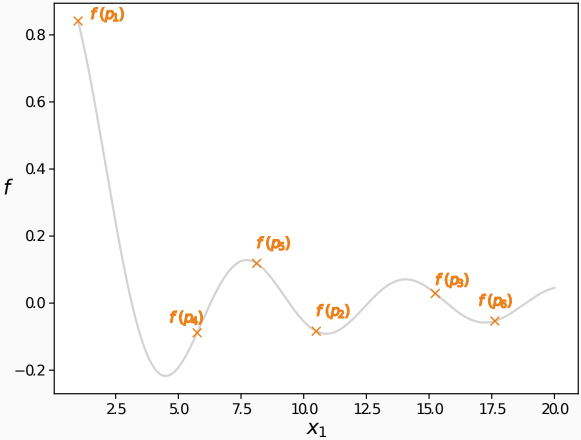


Figure 2: Sampling points on the surface found by hitting the map $f : \mathbb{R}^1 \to \mathbb{R}$

A brief one-dimensional motivation iv

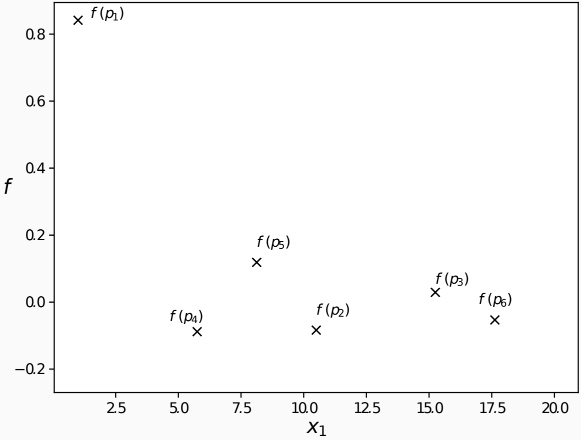


Figure 3: The information available to an algorithm (not very clear!)

A brief one-dimensional motivation v

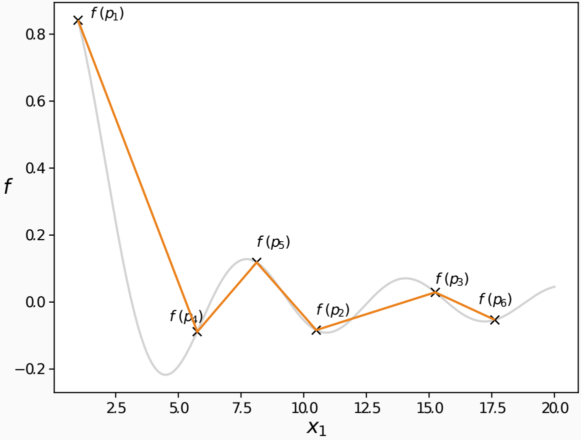


Figure 4: (Incomplete) geometric information found by building edges

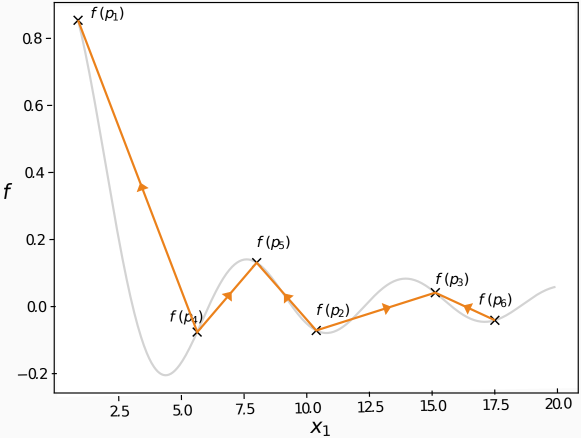


Figure 5: Directing the edges deduces even more information

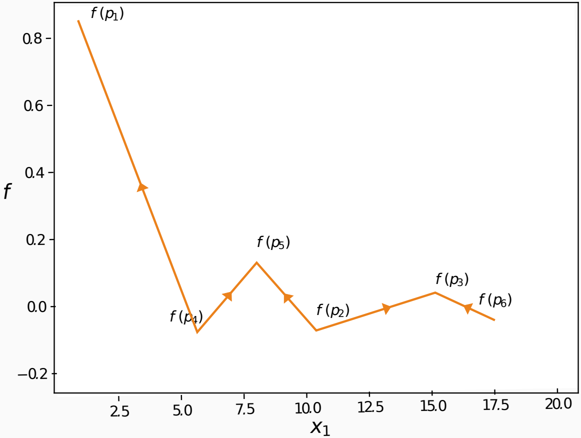


Figure 6: This geometric structure leaves us with a clearer picture

A brief one-dimensional motivation viii

- The number of local minima is at least 3 (by the mean value theorem)
- If we had just one fewer sampling point it would be impossible to deduce that there are 3 local minima
- On the other had if we had many more sampling points the number of minimisers would still only be 3 (a geometric **invariance!**)
- We want an idea of how many sampling points we need to find all solutions
- We would also like to know if these solutions are close together or far apart etc.
- We want to identify regions where it is proven we will find solutions (**locally convex sub-domains** that can be used in local-minimisation)
- Finally we want to extend these ideas to higher dimensions

Onward to the second dimension!

Example

Consider a more complex optimisation problem in two dimensions

$$\min f, \; x \in [0, 10] \times [0, 10]$$

where

$$f(\mathbf{x}) = -\sin(2x_1 - 0.5\pi) - 3\cos(x_2) - 0.5x_1$$

Subject to the following non-linear constraints:

$$\begin{aligned}
(x_1 - 5)^2 + (x_2 - 5)^2 + 5\sqrt{x_1 x_2} - 29 &\geq 0 \\
(x_1 - 6)^4 - x_2 + 2 &\geq 0 \\
9 - x_2 &\geq 0
\end{aligned}$$

2-dimensional surfaces ii

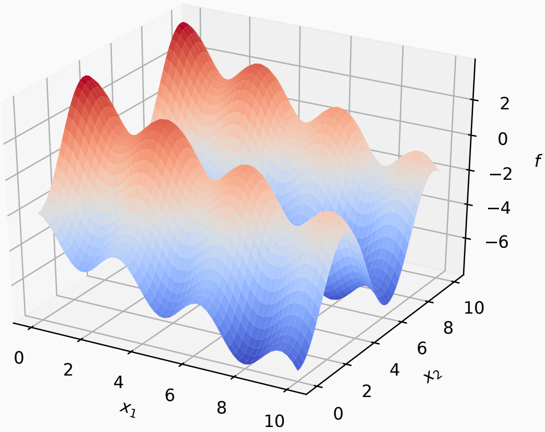


Figure 7: 3-dimensional surface plot of the example objective function

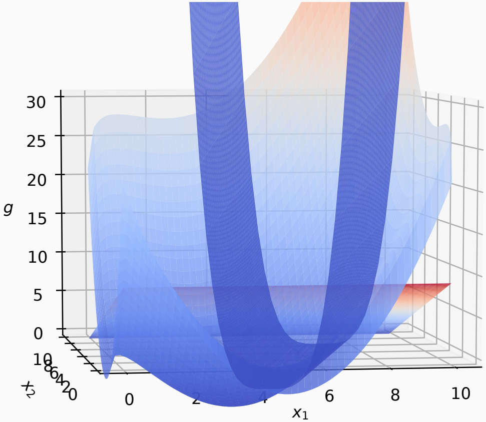


Figure 8: 3-dimensional surface plot of the example constraint functions

2-dimensional surfaces iv

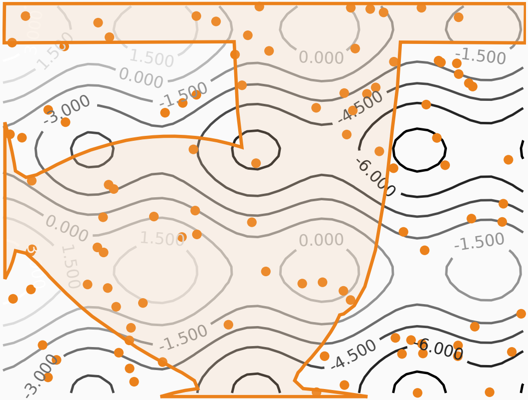


Figure 9: Contour plot of the problem, the shaded region violates the constraints, a set of random sampling points has been plotted on the surface

2-dimensional surfaces v

Many challenges are apparent:

- Already we can no longer use the simple graph structures from the 1-dimensional example since they do not cover the entire volume of 2-D space between points (also known as vertices in graph theory)
- Curse of dimensionality: when the dimensionality increases, the volume of the space increases so fast ($\mathcal{O}(2^d)$) that the available geometric data become sparse for the same number of sampling points
- Intuitively most algorithms utilising some kind multi-start routine will pass several sampling points that lead to the same solution several times

Understanding the problem

Understanding the problem i

- We can no longer track the invariant geometric features since the sampling points do not connect in such a way that it covers the full volume of space between points in the same way as the one-dimensional case
- We no longer have rigorous proofs of regions containing solutions (locally convex sub-domains)
- We can keep "guessing" and using multi-start routines, but we would potentially need thousands of sampling points every time even on very simple problems to cover the vast volume of the search (hyper-)space
- In addition, many local minimisations will be wasted only to find the same solution, this problem is exacerbated in even higher dimensions

Understanding problem means understanding hyperspace

What do we know about hyperspace? How do we use this knowledge?

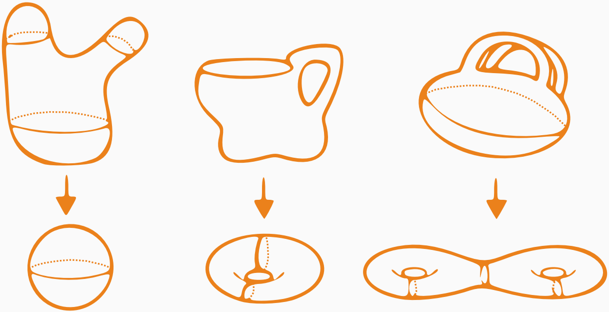
Understanding hyperspace through algebraic topology i

- Topology is the study of properties of geometric objects that endure when the objects are subjected to continuous transformations ("rubber sheet geometry")
- Many of these properties are readily extendable to arbitrarily high-dimensions
- In order to compute these properties, we need something to count (an algebra!)
- The field of algebraic topology studies the various ways in which we can use abstract algebra to study topological spaces

Topology is preserved under continuous transformations Link

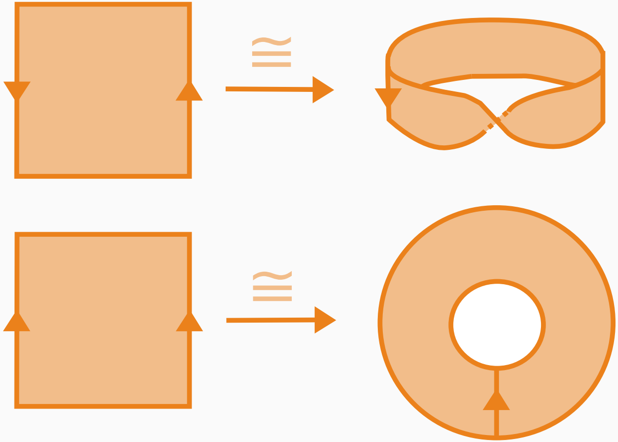
Topology is preserved under continuous transformations

Many surfaces have homeomorphic topological properties:

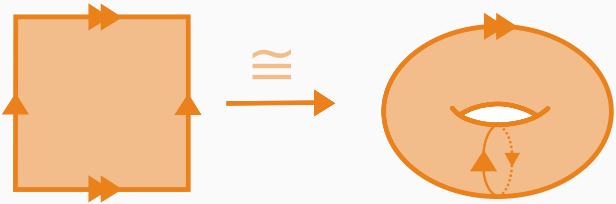


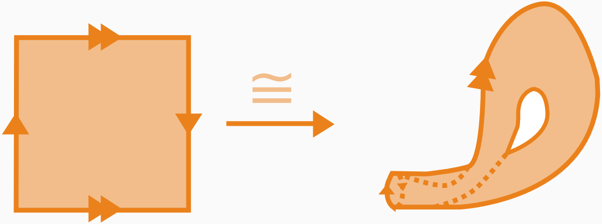
In 2-dimensional space the classification theorem proves that all possible closed or bounded surfaces are homeomorphic to the sphere, or a connected sum of tori or a connected sum of projective planes.

The Möbius band and the annulus



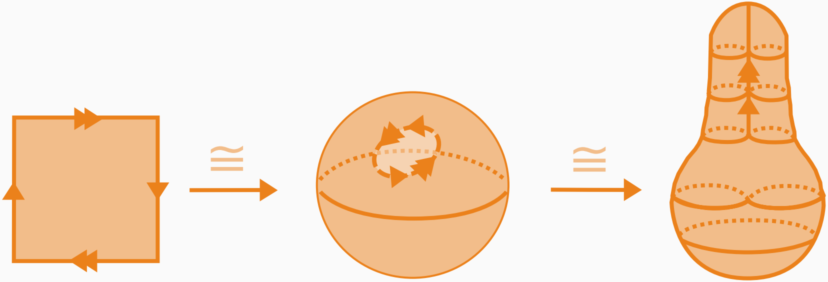
The torus and the Klein bottle





Topological surfaces and their plane models iii

The real projective plane



Nomenclature for developing a simplicial homology

In the development of shgo we require several concepts from algebraic and combinatorial topology [Hatcher, 2002, Henle, 1979]. We will start with the basic building blocks of a simplicial complex:

Definition

A **k-simplex** is a set of $n+1$ vertices in a convex polyhedron of dimension n. Formally if the $n+1$ points are the $n+1$ standard $n+1$ basis vectors for $\mathbb{R}^{(n+1)}$. Then the n-dimensional k-simplex is the set

$$S^n = \left\{ (t_1, \ldots, t_{n+1}) \in \mathbb{R}^{n+1} \mid \sum_{1}^{n+1} t_{n+1} = 1, t_i \geq 0 \right\}$$

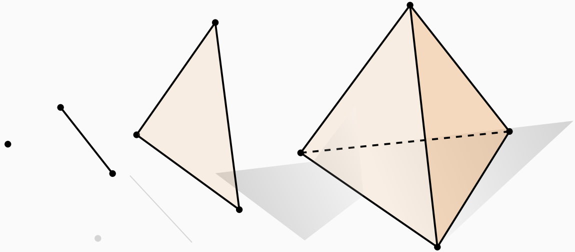


Figure 10: A 0-simplex (point), 1-simplex (edge), 2-simplex (triangle) and a 3-simplex (tetrahedron) (Figure adapted from [Keenan Crane, 2013])

Nomenclature for developing a simplicial homology iii

Definition

A **simplicial complex** $\mathcal{H}$ is a set $\mathcal{H}^0$ of vertices together with sets $\mathcal{H}^n$ of n-simplices, which are $(n+1)$-element subsets of $\mathcal{H}^0$. The only requirement is that each $(k+1)$-elements subset of the vertices of an n-simplex in $\mathcal{H}^n$ is a k-simplex, in $\mathcal{H}^k$.

Definition

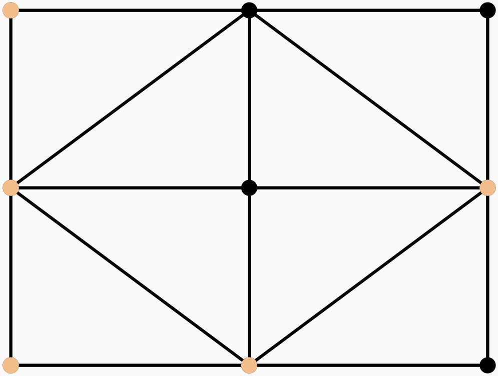
A **k-chain** is a union of simplices.

Examples:

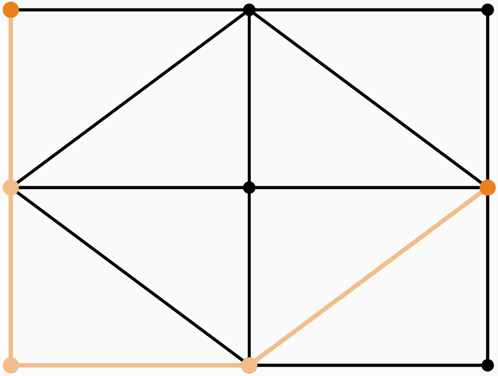
0-chain	**1-chain**	**2-chain**
A union of vertices.	A union of edges.	A union of triangles.

Nomenclature for developing a simplicial homology v

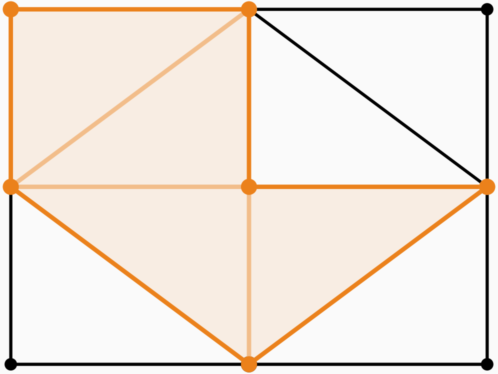
A 0-chain:



A 1-chain:

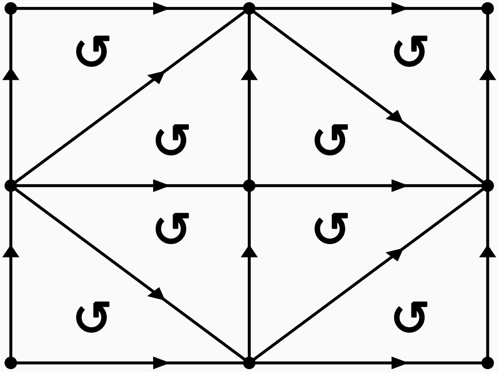


A 2-chain:

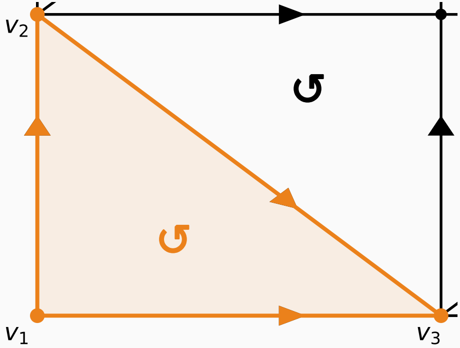


- $C(\mathcal{H}^k)$ denotes a $k-$chain of $k-$simplices.
- A vertex in $\mathcal{H}^0$ is denoted by v_i.
- If v_i and v_j are two endpoints of a directed 1-simplex in $\mathcal{H}^1$ from v_i to v_j then the symbol $\overline{v_i v_j}$ represents the 1-simplex
- This 1-simplex is bounded by the $0-$chain $\partial\left(\overline{v_i v_j}\right) = v_j - v_i$
- A 2-simplex consisting of three vertices v_i, v_j and v_k directed as $\overline{v_i v_j v_k}$ has the boundary of directed edges $\partial\left(\overline{v_i v_j v_j}\right) = \overline{v_i v_j} + \overline{v_j v_k} + \overline{v_j v_i}$.

A directed simplicial complex allows us to build an integral homology:

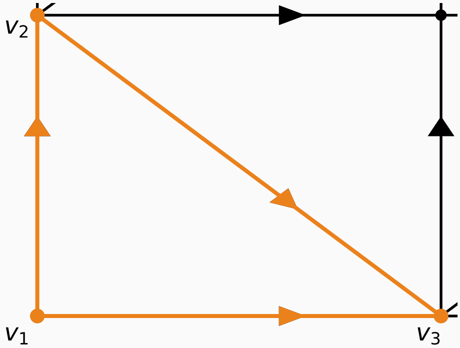


A directed 2-simplex in the directed simplicial complex



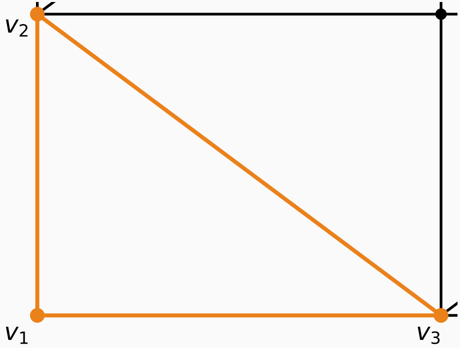
Nomenclature for developing a simplicial homology xi

The boundary operator acting on a directed simplex the edges of the directed 2-simplex: $\partial\left(\overline{v_1 v_2 v_3}\right) = \overline{v_1 v_3} - \overline{v_3 v_2} - \overline{v_2 v_1}$.



Note that in the **mod 2** homology the 1-chain $\overline{v_1 v_3} + \overline{v_3 v_2} + \overline{v_2 v_1}$ forms a **cycle** and that

$\partial(\overline{v_1 v_3} + \overline{v_3 v_2} + \overline{v_2 v_1}) = (v_3 - v_1) + (v_2 - v_3) + (v_1 - v_2) = \emptyset$



N.B.

In the directed integral homology we have $\partial\left(\overline{v_1 v_3} - \overline{v_3 v_2} - \overline{v_2 v_1}\right) = (v_3 - v_1) - (v_2 - v_3) - (v_1 - v_2)$ which contains additional information about the path.

This is just one example of the trade off between computational complexity and the information retained when using a mod 2 homology vs. a directed integral homology. For example mod 2 homologies fail to distinguish non-orientable surfaces from orientable (ex. klein bottle is non-orientable while a torus is orientable, but they have the same algebraic groups in a mod 2 homology).

In this study we will utilise both these homologies.

Example

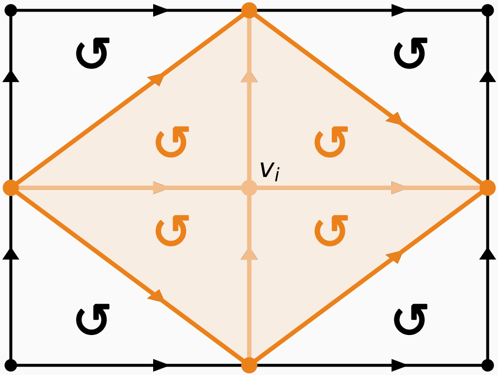
The directed simplicial complex on slide 21 is homologous to a torus. The chain complex has a non-zero 2-cycle by chaining all the 2-simplices $\partial\left(\sum_i^8 \mathcal{H}_i^2\right) = 0$. The Klein bottle has no such cycle.

Definition

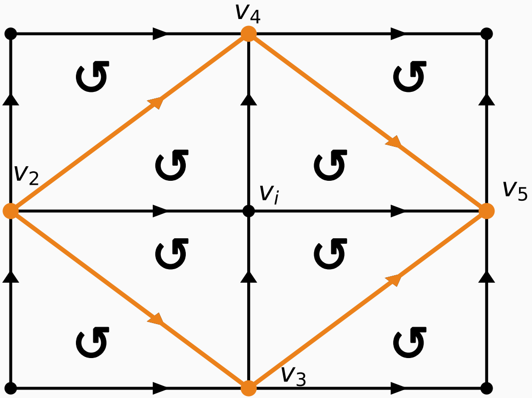
The **star** of a vertex v_i, written $\mathrm{st}\,(v_i)$, is the set of points Q such that every simplex containing Q contains v_i.

The $k-$chain $C(\mathcal{H}^k), k = n+1$ of simplices in $\mathrm{st}\,(v_i)$ forms a boundary cycle $\partial(C(\mathcal{H}^{n+1}))$ with $\partial\left(\partial(C(\mathcal{H}^{n+1}))\right) = \emptyset$. The faces of $\partial(\mathcal{H}^{n+1})$ are the bounds of the domain defined by $\mathrm{st}\,(v_i)$.

The domain defined by $\mathrm{st}(v_i)$:



The boundary $\partial\left(\mathrm{st}\left(v_i\right)\right) = \overline{v_2 v_3} + \overline{v_3 v_5} - \overline{v_5 v_4} - \overline{v_4 v_2}$:



Applying the simplicial homology

- Use simplicial complexes to extract information about the objective function (hyper-)surface using:
  - Simplicial integral homology theory
  - Discrete exterior calculus
  - Combinatorial and algebraic topology
- Algebraic topology theory is applied to provide rigorous convergence properties and higher performance properties
- To our knowledge, shgo is the first optimisation algorithm to make use of a homology theory (an algebraic topology theory about invariant geometric structures)
- Homology groups computed from sampling points on the hypersurface of objective functions allow us to deduce geometric features of the hypersurface that we can't visualize (a hypersurface has a dimension higher than 3)

Simplicial homology global optimisation

shgo: summary i

The algorithm itself consists of four major steps which will be described in detail:

1. Uniform sampling point generation of N vertices in the search space within the bounded and constrained subspace of Ω from which the $0-$chains of $\mathcal{H}^0$ are constructed
2. Construction of the directed simplicial complex $\mathcal{H}$ by triangulation of the vertices $h : \mathcal{P} \rightarrow \mathcal{H}$
3. Construction of the minimiser pool $\mathcal{M} \subset \mathcal{H}^0$ by repeated application of Sperner's lemma
4. Local minimisation using the starting points defined in $\mathcal{M}$

Computing the homology groups of hypersurfaces

How do we compute the homology group of an optimisation problem?

Overview: from Lipschitz surfaces to homology groups and the solution(s) of optimisation problems

$$
\begin{tikzcd}
 & & \mathbf{H}_i(\mathcal{H}), \mathbf{H}_i(\mathcal{S}) \arrow[r, "\text{Eilenberg-Steenrod Axioms}^{[1]}"] \arrow[d, leftrightarrow, "\forall i (\mathbf{H}_i(\mathcal{H}) \cong \mathbf{H}_i(\mathcal{S}))"] & \text{Invariance} \arrow[d, "\text{Finite cardinality}"] \arrow[dl, leftrightarrow] \\
\mathrm{f} \arrow[r, "h \circ f : \mathcal{P} \to \mathcal{H}"] & \mathcal{H} \arrow[ur, leftrightarrow, "\text{Simplicial homology}"] \arrow[r, leftrightarrow, "\mathcal{H} \cong \mathcal{K}", "k : \mathcal{K} \to \mathcal{S}"'] \arrow[dr, "\mathit{Sperner}"', "\partial C_i(\mathcal{H})"] & \mathcal{S} & f^*, \mathbf{x}^* \\
 & & \{\mathbf{x} | \forall x_i \leq \partial C(\mathcal{H}^k)(\mathbf{H}_i(\mathbf{x}) \geq 0)\} \arrow[r, "\text{DEC}^{[2]}", "\text{MVT}^{[3]}"'] & \nabla f(\mathbf{X}) = \bar{0} \arrow[u, "\text{Sets of solutions}"]
\end{tikzcd}
$$

1. [Eilenberg and Steenrod, 1952] , 2. Discrete exterior calculus , 3. (Discrete) Mean Value Theorem

Simplicial homology global optimisation: $h : \mathcal{P} \rightarrow \mathcal{H}$

shgo: $h : \mathcal{P} \rightarrow \mathcal{H}$ i

$$
\begin{tikzcd}
 & & \mathbf{H}_i(\mathcal{H}), \mathbf{H}_i(\mathcal{S}) \arrow[r, "\text{Eilenberg-Steenrod Axioms}"] \arrow[d, leftrightarrow, "\forall i(\mathbf{H}_i(\mathcal{H}) \cong \mathbf{H}_i(\mathcal{S}))"] & \text{Invariance} \arrow[d, "\text{Finite cardinality}"'] \arrow[dl, leftrightarrow] \\
f \arrow[r, "h \circ f : \mathcal{P} \rightarrow \mathcal{H}"] & \mathcal{H} \arrow[r, leftrightarrow, "\mathcal{H} \cong \mathcal{K}", "k : \mathcal{K} \rightarrow \mathcal{S}"'] \arrow[ur, leftrightarrow, "\text{Simplicial homology}"] \arrow[dr, "\text{Sperner}"', "\partial C_i(\mathcal{H})"] & \mathcal{S} & f^*, \mathbf{x}^* \\
 & & \{\mathbf{x} | \forall x_i \leq \partial C(\mathcal{H}^k)(\mathbf{H}_i(\mathbf{x}) \geq 0)\} \arrow[r, "\text{DEC}", "\text{MVT}"'] & \nabla f(\mathbf{X}) = \bar{0} \arrow[u, "\text{Sets of solutions}"]
\end{tikzcd}
$$

shgo: $h : \mathcal{P} \rightarrow \mathcal{H}$ ii

- We define the constructions used to build the simplicial complex on the hypersurface f from which we compute the homology groups
- $\mathcal{H}^0 := \mathcal{P}$ is the set of all vertices of $\mathcal{H}$ built from the set of feasible sampling points $\mathcal{P} = \{\mathbf{x} \in \mathcal{X} \mid \mathbf{g}(\mathbf{x}) \geq 0\}$
- The simplicial complex $\mathcal{H}$ is constructed by a triangulation connecting every vertex in $\mathcal{H}^0$
- The set $\mathcal{H}^1$ is constructed by directing every edge
- The edge is directed as $\overline{v_i v_j}$ from v_i to v_j iff $f(v_i) < f(v_j)$ so that $\partial\left(\overline{v_i v_j}\right) = v_j - v_i$
- Similarly an edge is directed as $\overline{v_j v_i}$ from v_j to v_i iff $f(v_i) > f(v_j)$ so that $\partial\left(\overline{v_j v_i}\right) = v_i - v_j$
- We let the higher dimensional simplices of $\mathcal{H}^k, k = 2, 3, \ldots n + 1$ be directed in any arbitrary direction which completes the construction of the complex $h : \mathcal{P} \rightarrow \mathcal{H}$

shgo: $h : \mathcal{P} \rightarrow \mathcal{H}$ iii

We can now use $\mathcal{H}$ to find the minimiser pool for the local minimisation starting points used by the algorithm:

Definition

A vertex v_i is a minimiser iff every edge connected to v_i is directed away from v_i, that is $\partial\left(\overline{v_i v_j}\right) = \left(v_{j \neq i} - v_i\right) \vee 0 \; \forall v_{j \neq i} \in \mathcal{H}^0$. The minimiser pool $\mathcal{M}$ is the set of all minimisers.

Example

The Ursem01 function for two dimensions is defined as follows [Gavana, 2016]

$$\min f, \; x \in \Omega = [0, 9] \times [-2.5, 2.5]$$

$$f(\mathbf{x}) = -\sin\left(2x_1 - 0.5\pi\right) - 3\cos\left(x_2\right) - 0.5x_1$$

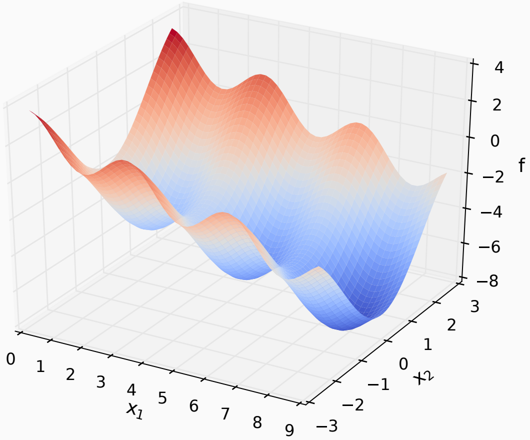


Figure 11: 3-dimensional plot of the Ursem01 function

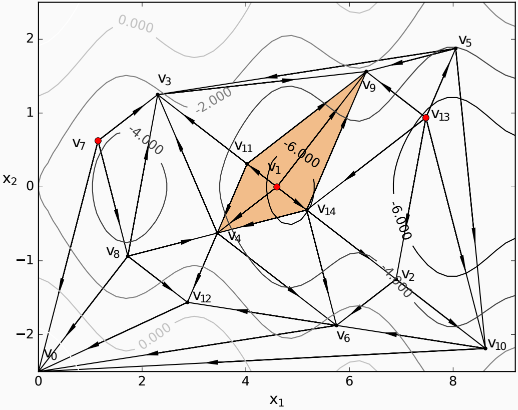


Figure 12: A directed complex $\mathcal{H}$ forming a simplicial approximation of f, three minimiser vertices $\mathcal{M} = \{v_1, v_7, v_{13}\}$ and the shaded domain $\mathrm{st}\,(v_1)$

Simplicial homology global optimisation: locally convex sub-domains

shgo: locally convex sub-domains i

$$
\begin{tikzcd}
 & & \mathbf{H}_i(\mathcal{H}), \mathbf{H}_i(\mathcal{S}) \arrow[r, "\text{Eilenberg-Steenrod Axioms}"] \arrow[d, leftrightarrow, "\forall i(\mathbf{H}_i(\mathcal{H})\cong\mathbf{H}_i(\mathcal{S}))"] & \text{Invariance} \arrow[d, "\text{Finite cardinality}"'] \arrow[dl, leftrightarrow] \\
\text{f} \arrow[r, "h\circ f:\mathcal{P}\to\mathcal{H}"] & \mathcal{H} \arrow[r, leftrightarrow, "\mathcal{H}\cong\mathcal{K}", "k:\mathcal{K}\to\mathcal{S}"'] \arrow[ur, leftrightarrow, "\text{Simplicial homology}"] \arrow[dr, "\text{Sperner} \quad \partial C_i(\mathcal{H})"'] & \mathcal{S} & f^*, \mathbf{x}^* \\
 & & \{\mathbf{x} | \forall x_i \leq \partial C(\mathcal{H}^k)(\mathbf{H}_i(\mathbf{x}) \geq 0)\} \arrow[r, "\text{DEC}", "\text{MVT}"'] & \nabla f(\mathbf{X}) = \bar{0} \arrow[u, "\text{Sets of solutions}"']
\end{tikzcd}
$$

shgo: locally convex sub-domains ii

- We want to find all the solutions of the problem
- The shgo algorithm finds sub-domains wherein a stationary point is guaranteed to be found
- Both these starting points and their domains allow us to find accurate solutions more easily

Theorem

*(**Stationary point in a minimiser star domain**) Given a minimiser $v_i \in \mathcal{M} \subseteq \mathcal{H}^0$ on the surface of a continuous objective function f with a compact bounded domain in $\mathbb{R}^n$ and range $\mathbb{R}$, there exists at least one stationary point of f within the domain defined by $st(v_i)$.*

shgo: locally convex sub-domains iii

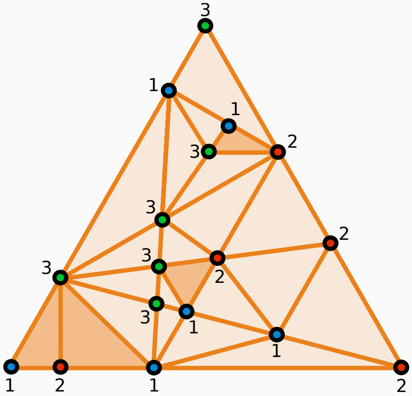
Overview:

- Find simplices with Sperner labels where each label represents a different $n+1$ label in every vector direction of the gradient vector field ∇f of f
- Of the $n+1$ Cartesian directions we require only a vector pointing towards a section defined by $n+1$ hyperplane cuts
- In a sense we extend the classical Brouwer's fixed point theorem [Brouwer, 1911] found in for example [Henle, 1979, p. 40] to optimisation problems with arbitrary constraints

Theorem

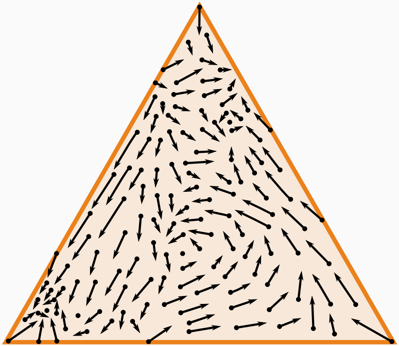
*(**Sperner's lemma [Sperner, 1928]**) Every Sperner labelling of a triangulation of a n-dimensional simplex contains a cell labelled with a complete set of labels: 1,2, . . . , n+1.*

A Sperner labelling, every vertex of the n-simplex is labelled with a set of labels $1, 2, \ldots, n+1$. Any vertices on the boundary $(n-1)$-simplices of the n-simplex may only contain the labels of its boundary vertices

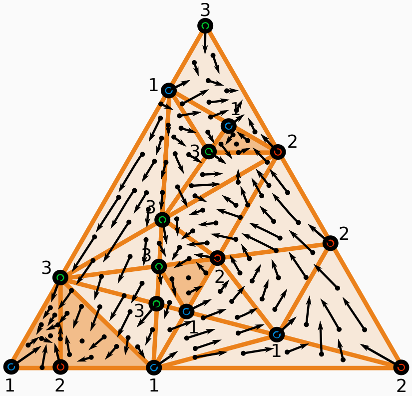


- The edge $\overline{13}$ may only contain vertices labelled either 1 or 3
- The edge $\overline{12}$ may only contain vertices labelled either 1 or 2
- The remainer of vertices inside the sub-triangulation may receive any arbitrary label in the set $1, 2, \ldots, n+1$

For example consider a vector field within a simplex. We may be interested in finding critical points where the vector field is stationary $V(P) = 0$ as in the proof of Brouwer's fixed point theorem:



We can devide the directions and assign a label to each of the vertices. Sperner's lemma gaurantees that there will be at least one sub-triangulation with the full set of labels:

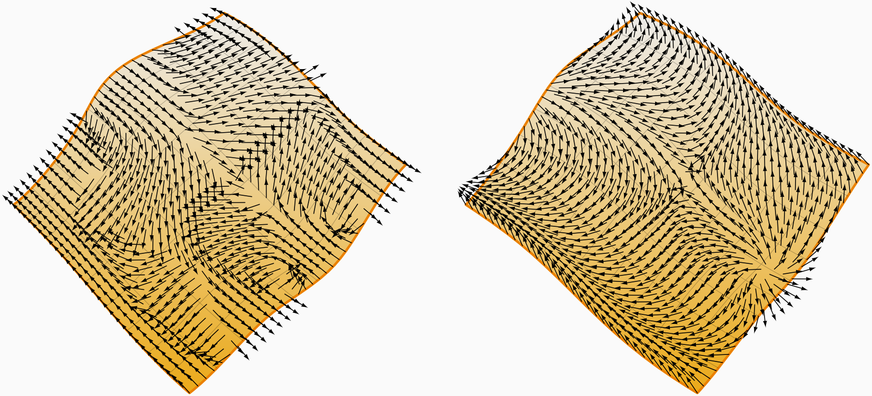


Example

It is proven that any simplex with a Sperner labelling must contain a sub-triangulation with another simplex that contains a Sperner labelling. Start by assigning every possible vector direction to a label. Then a simplex from the sub-triangulation must contain another sub-triangulation containing a Sperner simplex and so on until the sequence of sub-simplices produce a critical point.

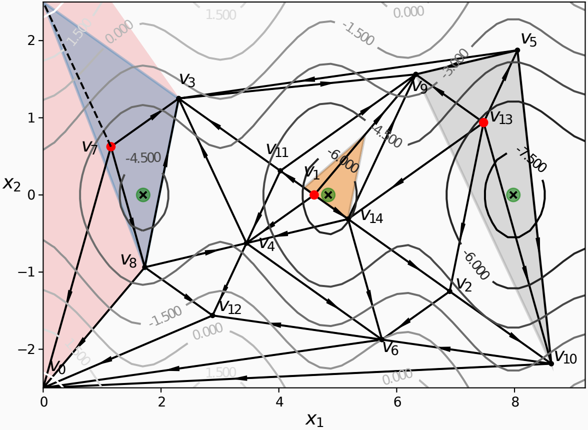
Brouwer used as a practical example in 3-dimensional space the fluid vector field of a coffee. **No matter how vigorously you stir your coffee, it is proven there is at least one point where the coffee remains stationary at any given time.**

On any gradient vector field, we can find sub-divisions containing Sperner simplices by sampling the surface (figure adapted from Rhino docs ‣ Link)



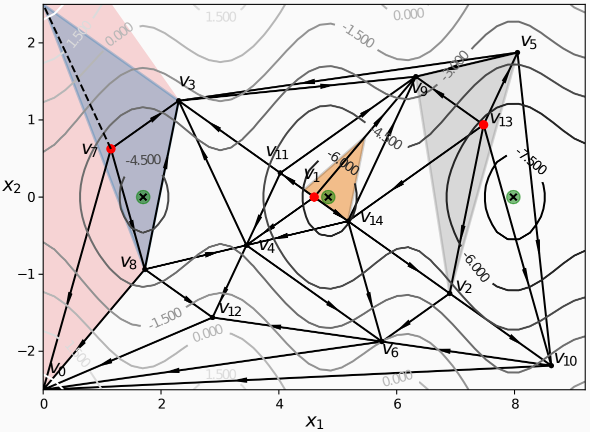
shgo: locally convex sub-domains x

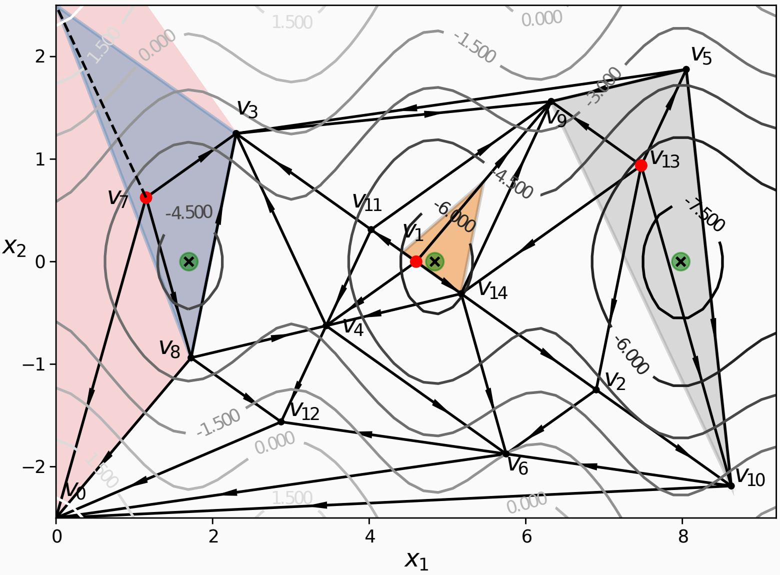
Possible Sperner simplices around domain v_7, domain v_1 and v_{13}



shgo: locally convex sub-domains xi

The domain $\partial(v_{13})$ cannot be further refined by the theorem





v_5
v_3
v_9
v_{13}
v_7
v_{11}
v_1
v_{14}
v_4
v_8
v_2
v_{12}
v_6
v_{10}
v_0
x_1
x_2
1.500
0.000
-1.500
-3.000
-4.500
-6.000
-7.500
0
2
4
6
8
2
1
0
−1
−2

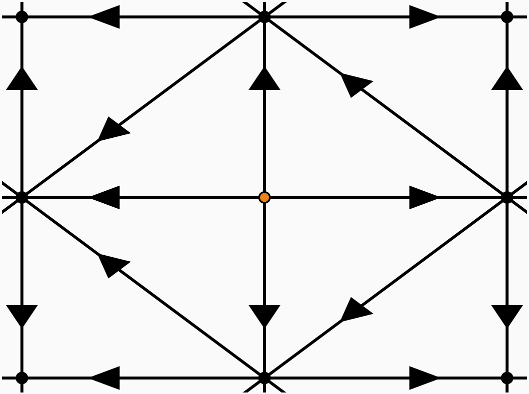
Simplicial homology global optimisation: invariance

$$\begin{tikzcd}
 & & \mathbf{H}_i(\mathcal{H}), \mathbf{H}_i(\mathcal{S}) \arrow[r, "\text{Eilenberg-Steenrod Axioms}"] \arrow[dl, leftrightarrow, "\text{Simplicial homology}"'] \arrow[d, leftrightarrow, "\forall i (\mathbf{H}_i(\mathcal{H}) \cong \mathbf{H}_i(\mathcal{S}))"] & \text{Invariance} \arrow[d, "\text{Finite cardinality}"] \arrow[dl, leftrightarrow] \\
\mathrm{f} \arrow[r, "h \circ f : \mathcal{P} \to \mathcal{H}"] & \mathcal{H} \arrow[r, leftrightarrow, "\mathcal{H} \cong \mathcal{K}", "k : \mathcal{K} \to \mathcal{S}"'] \arrow[dr, "\textit{Sperner} \quad \partial C_i(\mathcal{H})"] & \mathcal{S} & f^*, \mathbf{x}^* \\
 & & \{\mathbf{x} | \forall x_i \leq \partial C(\mathcal{H}^k)(\mathbf{H}_i(\mathbf{x}) \geq 0)\} \arrow[r, "\text{DEC}", "\text{MVT}"'] & \nabla f(\mathbf{X}) = \bar{0} \arrow[u, "\text{Sets of solutions}"']
\end{tikzcd}$$

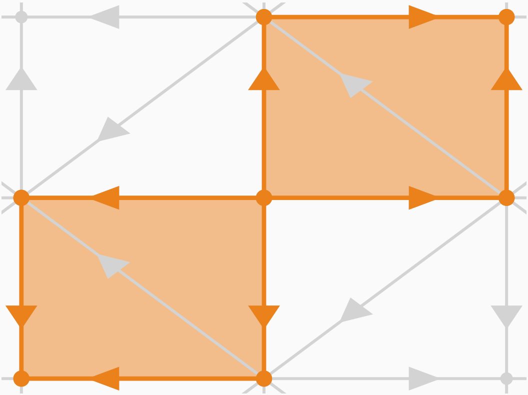
shgo: invariance ii

- For black box functions there is no way to know if the number and distribution of sampling points is adequate to find all the solutions without more information (for example if the number of local minima are known in the problem)
- However, we would still like to ensure that we don't "over sample" too much or waste time finding the same solution to the problem (all of which cost computational resources)
- First, the compact invariance theorem proves that this never happens in a compact space and in addition the algorithm converges to all solutions of the problem
- The proof relies on a homomorphism between the simplicial complex $\mathcal{H}$ constructed in a compact space and the homology (mod 2) groups of a constructed surface $\mathcal{S}_g$ and its triangulation $\mathcal{K}$ (with $\mathbf{H}_k(\mathcal{K}) \cong \mathbf{H}_k(\mathcal{S})\ \forall k \subset \mathbb{Z}$) on its surface on which we can invoke the invariance theorem

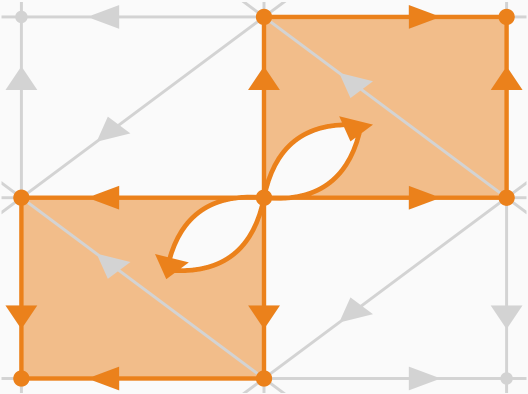
Construction of $\mathcal{S}_g$: Start by identifying a minimizer point in the $\mathcal{H}^1$ ($\cong$ $\mathcal{K}^1$) graph



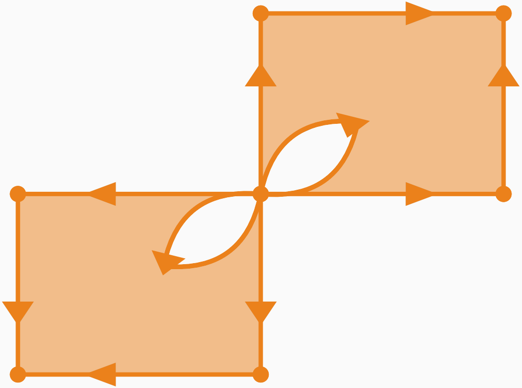
By construction, our initial complex exists on the (hyper-)surface of an n-dimensional torus $\mathcal{S}_0$ such that the rest of $\mathcal{K}^1$ is connected and compact



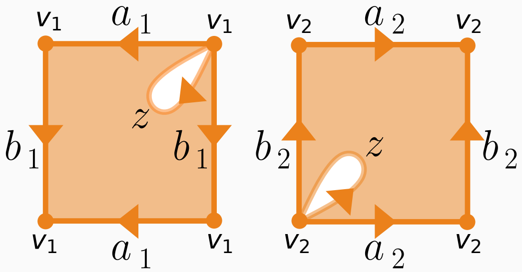
We puncture a hypersphere at the minimiser point and identify the resulting edges (or $(n-1)$-simplices in higher dimensional problems)



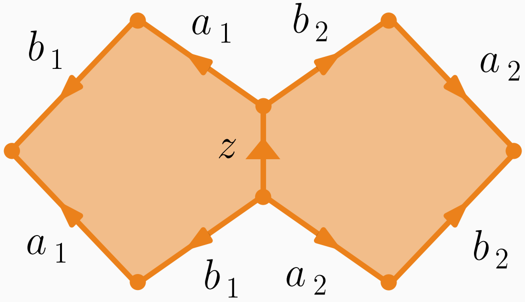
Shrink (*a topoligical (ie continuous) transformation*) the remainder of the simplicial complex to the faces and vertices of our (hyper-)plane model



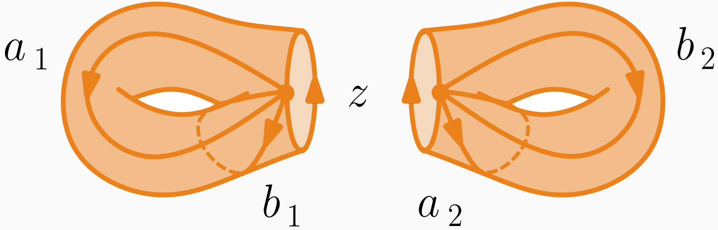
Make the appropriate identifications for $\mathcal{S}_0$ and $\mathcal{S}_1$



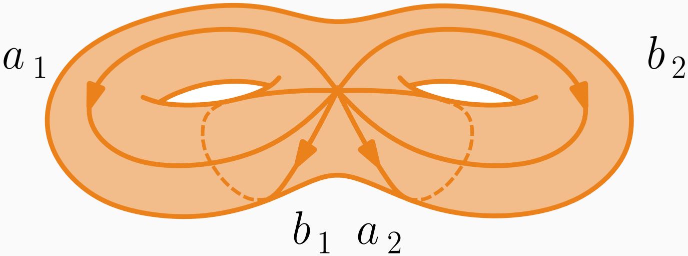
Glue the indentified and connected face z (a $(n-1)$-simplex) that resulted from the hypersphere puncture



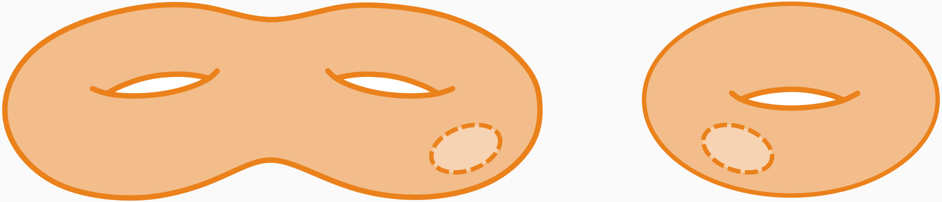
The other faces (ie $(n-1)$-simplices) are connected in the usual way for tori constructions)



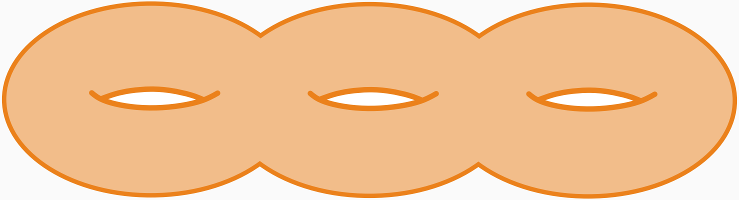
The resulting (hyper-)surface $\mathcal{S} = \mathcal{S}_0 \,\#\, \mathcal{S}_1$



We can repeat the process with $\mathcal{S}_0 \# \mathcal{S}_1$ for a new minimiser point and corresponding hypersurface $\mathcal{S}_2$ without loss of generality

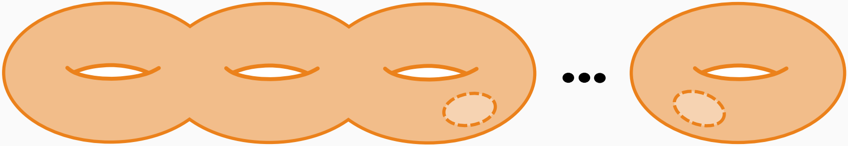


$\mathcal{S} = \mathcal{S}_0 \# \mathcal{S}_1 \# \mathcal{S}_2$



Repeat this process for every minimiser point in the set $\mathcal{M}$

$$\mathcal{S}_g := \mathcal{S}_0 \# \mathcal{S}_1 \# \cdots \# \mathcal{S}_{g-1} \qquad (g \text{ times})$$



- In homology thoery a theorem known as the Invariance Theorem can be extended to higher dimensional triangulable spaces using singular homology through the famous Eilenberg-Steenrod Axioms [Eilenberg and Steenrod, 1952, Henle, 1979]
- As a direct consequence any triangulation of $\mathcal{S}_g$ will produce the same homology groups for $\mathcal{K}$
- Adding any new sampling point will produce the same homology groups since $\text{rank}(\mathbf{H}_1(\mathcal{K}))$ (the "number of holes in $\mathcal{S}_g$") remains unchanged and it is thus not possible for a new vertex to be wrongly identified as a minimiser in the triangulation $\mathcal{H}$

N.B.

Any further refinement in the simplicial complex by further sampling does not increase the number of locally convex sub-domains in a compact space!

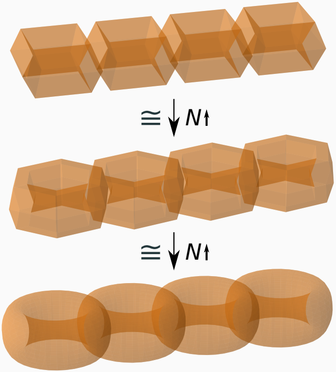


Figure 13: Refining the simplicial complex $\mathcal{K}$ built on the connected g sum of g tori $\mathcal{S}_g$ does not change the Betti numbers of the surface (also related to the Euler characteristic)

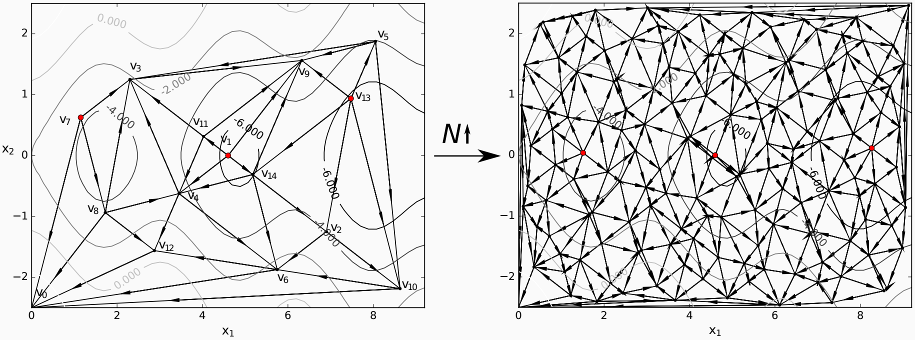


Figure 14: Further refinement of the simplicial complex from the example problem doesn't increase the number of locally convex sub-domains extracted by shgo because of the homomorphims between the homology groups of $\mathcal{H}$ and $\mathcal{K}$

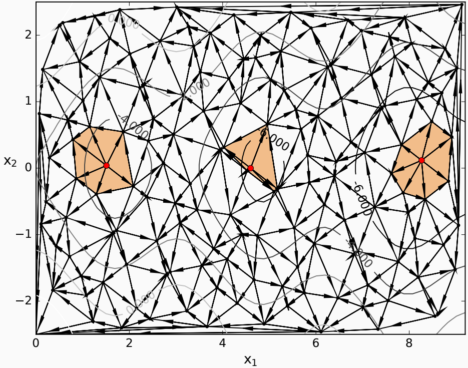


Figure 15: After increasing the number of sampling points the number of locally convex sub-domains from the example problem are still 3, however, the boundaries of the star domains have been further refined

shgo: invariance xviii

- shgo is proven to have a **stronger invariance** and **convergence** in the case where the constraints $\mathbf{g}$ are non-linear
- In addition we allow the objective function f to be non-continuous and non-linear

shgo: invariance xix

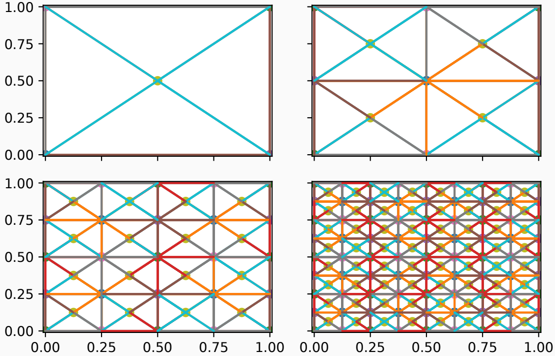


Figure 16: Simplicial sampling by sub-triangulation of hyper-rectangles

Example

We expand the bounds of the Ursem01 function for two dimensions [Gavana, 2016]

$$\min f, \ x \in [0, 10] \times [0, 10]$$

Subject to the following non-linear constraints:

$$
\begin{aligned}
(x_1 - 5)^2 + (x_2 - 5)^2 + 5\sqrt{x_1 x_2} - 29 \geq 0 \\
(x_1 - 6)^4 - x_2 + 2 \geq 0 \\
9 - x_2 \geq 0
\end{aligned}
$$

$$f(\mathbf{x}) = -\sin(2x_1 - 0.5\pi) - 3\cos(x_2) - 0.5x_1$$

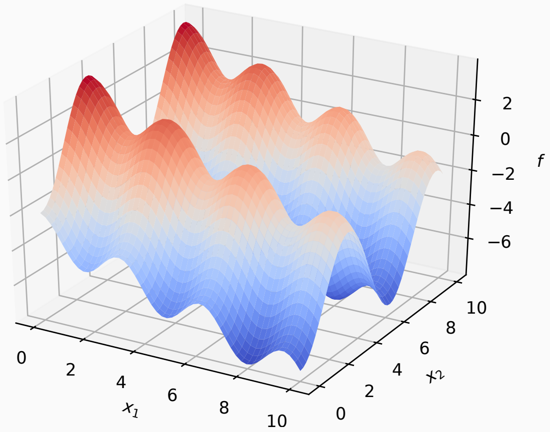
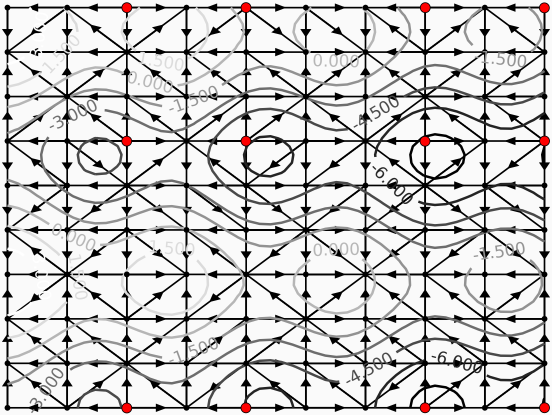


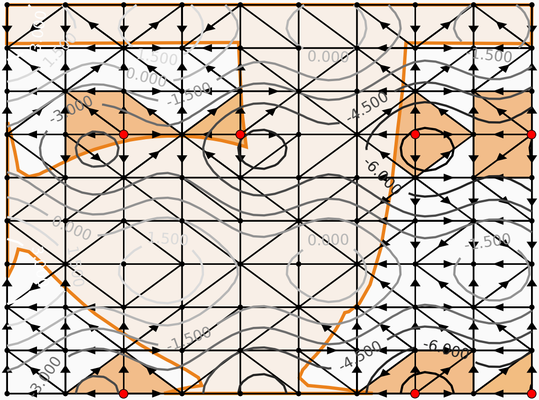
Figure 17: 3-dimensional plot of the Ursem01 function with expanded bounds

First consider $\mathcal{H}$ without the non-linear bounds, here $|\mathcal{M}| = 12$:

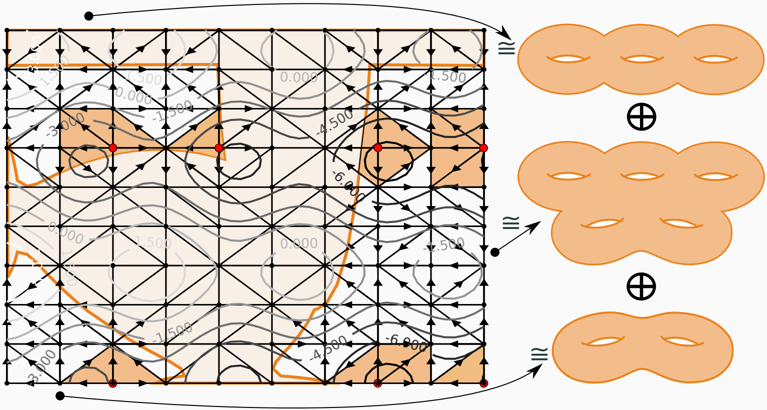


shgo: invariance xxiii

After applying the non-linear version of h, the non-linear bounds produce the following disconnected simplicial complexes:



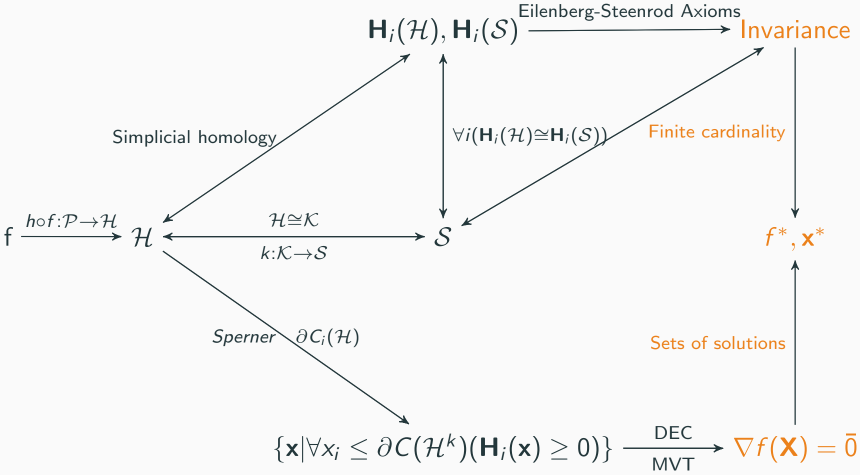
We use the fact that for abelian homology groups the rank is additive over arbitrary direct sums $\mathbf{rank}\left(\bigoplus_{i\in I}\mathbf{H_1}(\mathcal{K}_i)\right) = \sum_{i\in I}\mathbf{rank}(\mathbf{H_1}(\mathcal{K}_i))$:



But why?

Simplicial homology global optimisation: algorithm

shgo: algorithm i



shgo: algorithm ii

procedure INITIALISATION
  Input an objective function f, constraint functions $\mathbf{g}$ and variable bounds and $[\mathbf{l}, \mathbf{u}]^n$.
  Input N initial sampling points.
  Define a sampling sequence that generates a set $\mathcal{X}$ of sampling points in the unit hypercube space $[\mathbf{0}, \mathbf{1}]^n$
  Define the empty set $\mathcal{M}^E = \emptyset$ of vertices evaluated by a local minimisation.
end procedure
while TERM$(\mathbf{H}_1(\mathcal{H}), \min\{\mathcal{F}\})$ is False **do**
  procedure SAMPLING
    $\mathcal{P} = \emptyset$
    while $|\mathcal{P}| < N$ **do**
      Generate $N - |\mathcal{P}|$ sequential sampling points $\mathcal{X} \subset \mathbb{R}^n$
      Stretch $\mathcal{X}$ over the lower and upper bounds $[\mathbf{l}, \mathbf{u}]^n$

shgo: algorithm iii

$\mathcal{P} = \{\mathcal{X}_i \mid \mathbf{g}(\mathcal{X}_i) \geq 0, \forall \mathcal{X}_i \in \mathcal{X}\} \cup \mathcal{P}$ ▷ (Find $\mathcal{P}$ in the feasible subset Ω by discarding any points mapped outside the linear constraints g and adding to the current set of $\mathcal{P}$.)

Set $\mathcal{X} = \emptyset$

end while

Find $\mathcal{F}$ from the objective function $f : \mathcal{P} \to \mathcal{F}$ for any new points in $\mathcal{P}$

end procedure

procedure CONSTRUCT/APPEND DIRECTED COMPLEX $\mathcal{H}$

Calculate $\mathcal{H}$ from $h : \mathcal{P} \to \mathcal{H}$ ▷ (If $\mathcal{H}$ was already constructed new points in $\mathcal{P}$ are incorporated into the triangulation.)

Calculate $\mathbf{H}_1(\mathcal{H})$

end procedure

procedure CONSTRUCT $\mathcal{M}$

Find $\mathcal{M}$ from the definitions of h.

shgo: algorithm iv

24: **end procedure**

25: **procedure** LOCAL MINIMISATION

26: Calculate the approximate local minima of f using a local minimisation routine with the elements of $\mathcal{M} \setminus \mathcal{M}^E$ as starting points. ▷ Process the most promising points first.

27: $\mathcal{M}^E = \mathcal{M}^E \cap \mathcal{M}$ ▷ This excludes the evaluation any element $v_i \in \mathcal{M}$ that is known to be the only point that in the domain $\partial \mathrm{st}(v_j)$ where v_j is known to any point already used as a starting point in Step 27. If any new $v_i \in \mathcal{M}$ not in $\mathcal{M}^E$ is known to be the only point $\partial \mathrm{st}(v_j)$ it can also be excluded.

28: Add the function outputs of the local minimisation routine to $\mathcal{F}$

29: **end procedure**

30: Find new value of **TERM**$(\mathbf{H}_1)(\mathcal{H}, \min\{\mathcal{F}\})$

31: **end while**

shgo: algorithm v

procedure Process return objects
    Order the final outputs of the minima of f found in the local minimisation step to find the approximate global minimum.
end procedure

return the approximate global minimum and a list of all the minima found in the local minimisation step.

Properties

Properties of shgo:

- Convergence to a global minimum assured
- Allows for non-linear constraints in the problem statement
- Extracts all the minima in the limit of an adequately sampled search space (ie attempts to find all the (quasi-)equilibrium solutions)
- Progress can be tracked after every iteration through the calculated homology groups
- Competitive performance compared to state of the art black-box solvers
- All of the above properties hold for non-continuous functions with non-linear constraints assuming the search space contains any sub-spaces that are continuous and convex

Experimental results

Open-source black-box algorithms i

- Here we compare shgo with the following algorithms:
  - topographical global optimization (TGO) [Henderson et al., 2015]
  - basinhopping (BH) [Li and Scheraga, 1987, Wales, 2003, Wales and Doye, 1997, Wales and Scheraga, 1999]
  - differential evolution (DE) [Storn and Price, 1997]
- BH and DE are readily available in the SciPy project [Jones et al., 01]
- BH is commonly used in energy surface optimisations [Wales, 2015]
- DE has also been applied in optimising Gibbs free energy surfaces for phase equilibria calculations [Zhang and Rangaiah, 2011]
- SciPy global optimisation benchmarking test suite [Adorio and Dilman, 2005, Gavana, 2016, Jamil and Yang, 2013, Mishra, 2007, Mishra, 2006, NIST, 2016]

Open-source black-box algorithms ii

- The test suite contains multi-modal problems with box constraints, they are described in detail in `http://infinity77.net/global_optimization/` Link
- The stochastic algorithms (BH and DE) used the starting points provided by the test suite
- Stopping criteria $pe = 0.01\%$
- For every test the algorithm was terminated if the global minimum was not found after 10 minutes of processing time and the test was flagged as a fail
- For comparisons we used normalised performance profiles [Dolan and Moré, 2002] using function evaluations and processing time as performance criteria
- In total 180 test problems were used

Open-source black-box algorithms iii

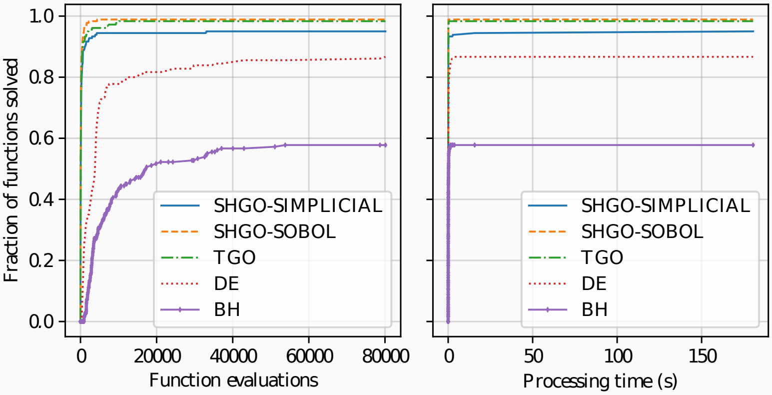


Figure 18: Performance profiles for SHGO, TGO, DE and BH

Open-source black-box algorithms iv

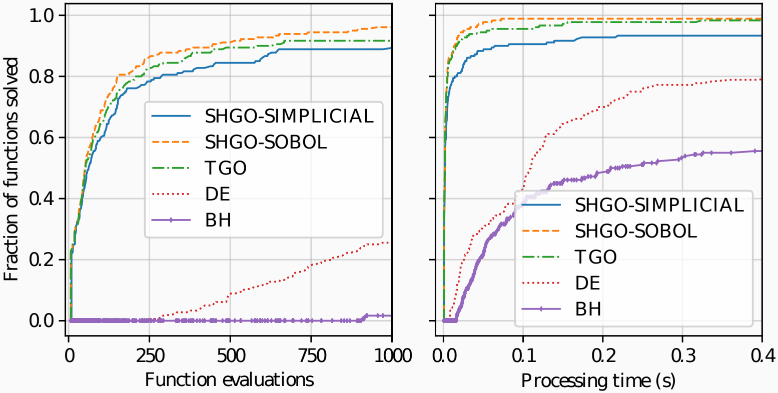


Figure 19: Performance profiles with ranges f.e. $= [0, 1000]$ and p.t. $= [0, 0.4]$

Open-source black-box algorithms v

- shgo-sobol was the best performing algorithm
- ... followed closely by tgo and shgo-simpl
- shgo-sobol tends to outperform tgo, solving more problems for a given number of function evaluations as expected for the same sampling point sequence
- tgo produced more than one starting point in the same locally convex domain while shgo is guaranteed to only produce one after adequate sampling
- While shgo-simpl has the advantage of having the theoretical guarantee of convergence, the sampling sequence has not been optimised yet requiring more function evaluations with every iteration than shgo-sobol

Linear-constrained optimisation problems i

- The DISIMPL algorithm was recently proposed by [Paulavičius and Žilinskas, 2014]
- The experimental investigation shows that the proposed simplicial algorithm gives very competitive results compared to the DIRECT algorithm [Paulavičius and Žilinskas, 2016]
- More recently the Lc-DISIMPL variant of the algorithm was developed to handle optimisation problems with linear constraints [Paulavičius and Žilinskas, 2016]
- Test on 22 optimisation problems again using the stopping criteria $pe = 0.01\%$
- Lc-DISIMPL-v, PSwarm (avg), DIRECT-L1 results produced by [Paulavičius and Žilinskas, 2016]

Linear-constrained optimisation problems ii

Table 1: Performance over all 22 test problems.

problem	algorithm	f.e.	runtime (s)
Average	SHGO-simplicial	65	0.012852
	SHGO-sobol	88	0.004144
	TGO	100	0.004542
	Lc-DISIMPL-v	366	-
	Lc-DISIMPL-c	$>$5877	-
	PSO (avg)	3011	-
	DIRECT-L1 (pp $=10$)	$>$17213	-
	DIRECT-L1 (pp $=10^2$)	$>$28421	-
	DIRECT-L1 (pp $=10^6$)	$>$75113	-

Table 2: Performance over all 22 test problems.

problem	algorithm	f.e.	nlmin	nulmin	runtime (s)
All	shgo-simpl	1463	26	26	0.27294
	shgo-sobol	1864	23	23	0.091168
	tgo	2123	29	25	0.093607

Linear-constrained optimisation problems iv

- The higher performance of shgo compared to tgo and DISIMPL is due to homological identification of unique locally convex sub-spaces
- shgo had
  - no wasted local minimisations unlike tgo because the locally convex sub-spaces are proven to be unique
  - no need for switching between a local and global step as in DISIMPL because the homology group rank growth tracks the global progress every iteration without requiring further refinement in sub-spaces
- For the full table of results see https://stefan-endres.github.io/shgo/files/table.pdf
  ▸ Link

Conclusions

Conclusions i

- The shgo algorithm shows promising properties and performance
- On test problems with linear constraints it was shown to provide competitive results to the TGO, Lc-DISIMPL, PSwarm and DIRECT-L1 algorithms
- On black-box problems it was shown to provide competitive results to the TGO, BH and DE algorithms
- The use of a simplicial complex provides access to a wealth of tools from combinatorial topology and the growing field of computational homology
- It is hoped that these will drive further extensions and development

Conclusions ii

- Due to the useful characterisations of objective function hypersurfaces provided by the homology groups of the simplicial complex, shgo allows an optimisation practitioner with a useful visual tool for understanding and efficiently solving higher dimensional black and grey box optimisation problems
- It is especially appropriate for computationally expensive black and grey box functions common in science and engineering
- In addition because the homology groups can be calculated as sampling progresses an optimisation practitioner can both visualise the extent of the optimisation problems multi-modality and use intelligent stopping criteria for the sampling stage

Thank you for your time.

References

References i

Adorio, E. P. and Dilman, U. P. (2005).
MVF - Multivariate Test Functions Library in C for Unconstrained Global Optimization.
`http://www.geocities.ws/eadorio/mvf.pdf` [Accessed: September 2016].

Brouwer, L. E. J. (1911).
Über Abbildung von Mannigfaltigkeiten.
Mathematische Annalen, 71(1):97–115.

Dolan, E. D. and Moré, J. J. (2002).
Benchmarking optimization software with performance profiles.
Mathematical Programming, 91(2):201–213.

Eilenberg, S. and Steenrod, N. (1952).
Foundations of algebraic topology.
Mathematical Reviews (MathSciNet): MR14: 398b Zentralblatt MATH, Princeton, 47.

Endres, S. (2016–).
SHGO: Python implementation of the simplicial homology global optimisation algorithm.
[Online; accessed 2016-11-04].

Gavana, A. (2016).
Global Optimization Benchmarks and AMPGO.
http://infinity77.net/global_optimization/index.html
[Accessed: September 2016].

References iii

Hatcher, A. (2002).
Algebraic topology.
Cambridge University Press, Cambridge.

Henderson, N., de Sá Rêgo, M., Sacco, W. F., and Rodrigues, R. A. (2015).
A new look at the topographical global optimization method and its application to the phase stability analysis of mixtures.
Chemical Engineering Science, 127:151–174.

Henle, M. (1979).
A Combinatorial Introduction to Topology.
Unabriged Dover (1994) republication of the edition published by WH Greeman & Company, San Francisco, 1979.

- Jamil, M. and Yang, X.-S. (2013).
  A Literature Survey of Benchmark Functions For Global Optimization Problems Citation details: Momin Jamil and Xin-She Yang, A literature survey of benchmark functions for global optimization problems.
  Int. Journal of Mathematical Modelling and Numerical Optimisation, 4(2):150–194.
- Jones, E., Oliphant, T., Peterson, P., et al. (2001–).
  SciPy: Open source scientific tools for Python.
  [Online; accessed 2016-11-04].
- Keenan Crane, Fernando de Goes, M. D. P. S. (2013).
  Digital geometry processing with discrete exterior calculus.
  In *ACM SIGGRAPH 2013 courses*, SIGGRAPH '13, New York, NY, USA. ACM.

Li, Z. and Scheraga, H. A. (1987).
Monte carlo-minimization approach to the multiple-minima problem in protein folding.
Proceedings of the National Academy of Sciences, 84(19):6611–6615.

Mishra, S. (2007).
Some new test functions for global optimization and performance of repulsive particle swarm method.
http://mpra.ub.uni-muenchen.de/2718/ [Accessed: September 2016].

- Mishra, S. K. (2006).
  Global Optimization by Differential Evolution and Particle Swarm Methods Evaluation on Some Benchmark Functions.
  http://dx.doi.org/10.2139/ssrn.933827 [Accessed: September 2016].
- NIST (2016).
  NIST StRD Nonlinear Regression Problems.
  http://www.itl.nist.gov/div898/strd/nls/nls_main.shtml [Accessed: September 2016].
- Paulavičius, R., Sergeyev, Y. D., Kvasov, D. E., and Žilinskas, J. (2014).
  Globally-biased disimpl algorithm for expensive global optimization.
  Journal of Global Optimization, 59(2):545–567.

References vii

Paulavičius, R. and Žilinskas, J. (2014).
Simplicial lipschitz optimization without the lipschitz constant.
Journal of Global Optimization, 59(1):23–40.

Paulavičius, R. and Žilinskas, J. (2016).
Advantages of simplicial partitioning for lipschitz optimization problems with linear constraints.
Optimization Letters, 10(2):237–246.

Sperner, E. (1928).
Neuer beweis für die invarianz der dimensionszahl und des gebietes.
Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg, 6(1):265.

- Storn, R. and Price, K. (1997).
  Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces.
  Journal of Global Optimization, 11(4):341–359.
- Wales, D. (2003).
  Energy landscapes: Applications to clusters, biomolecules and glasses.
  Cambridge University Press.
- Wales, D. J. (2015).
  Perspective: Insight into reaction coordinates and dynamics from the potential energy landscape.
  Journal of Chemical Physics, 142(13).

- Wales, D. J. and Doye, J. P. (1997).
  Global optimization by basin-hopping and the lowest energy structures of lennard-jones clusters containing up to 110 atoms.
  The Journal of Physical Chemistry A, 101(28):5111–5116.
- Wales, D. J. and Scheraga, H. A. (1999).
  Global optimization of clusters, crystals, and biomolecules.
  Science, 285(5432):1368–1372.
- Zhang, H. and Rangaiah, G. P. (2011).
  A Review on Global Optimization Methods for Phase Equilibrium Modeling and Calculations.
  The Open Thermodynamics Journal, pages 71–92.

Questions?

Backup slides: Overview of proof of the stationary point theorem i

Theorem

*(**Stationary point in a minimiser star domain**) Given a minimiser $v_i \in \mathcal{M} \subseteq \mathcal{H}^0$ on the surface of a continuous, Lipschitz smooth objective function f with a compact bounded domain in $\mathbb{R}^n$ and range $\mathbb{R}$. For any $n-$dimensional $k-$chain $C(\mathcal{H}^k), k = n+1$ with subset of edges $E \subseteq \{C(\mathcal{H}^k), k = n+1\} \subset \mathcal{H}^1$. If v_i has incidence on a set of edges E, then the chain of simplices containing E defines a $k-$chain $C(\mathcal{D}^k)$, $\mathcal{D}^k \subseteq \mathcal{H}^k$, $k = n+1$ near v_i with every vertex in $C(\mathcal{D}^k)$ connected to v_i. There exists at least one stationary point of f within the domain defined by the boundary cycle $\partial(\mathcal{D}^{n+1})$.*

Backup slides: Overview of proof of the stationary point theorem ii

Overview

- Find a simplex with a Sperner labelling where each label represents a different $n+1$ label in every vector direction of the gradient vector field ∇f of f
- Of the $n+1$ Cartesian directions we require only a vector pointing towards a section defined by $n+1$ hyperplane cuts
- The remainder of the proof then proceeds as usual for Brouwer's fixed point theorem [Brouwer, 1911] found in for example [Henle, 1979, p. 40] utilising Sperner's lemma

Backup slides: Overview of proof of the stationary point theorem iii

Theorem

(**Sperner's lemma [Sperner, 1928]**) *Every Sperner labelling of a triangulation of a n-dimensional simplex contains a cell labelled with a complete set of labels: 1,2, . . . , n+1.*

- For any minimiser $v_i \in \mathcal{M} \subseteq \mathcal{H}^0$ we have by construction that for any vertex v_j with incidence on a connecting edge $\overline{v_i v_j}$ that $f(v_i) < f(v_j)$
- By the MVT there is at least one point on $\overline{v_i v_j}$ where ∇f points towards a Cartesian direction in a section that can receive a unique Sperner label

Backup slides: Overview of proof of the stationary point theorem iv

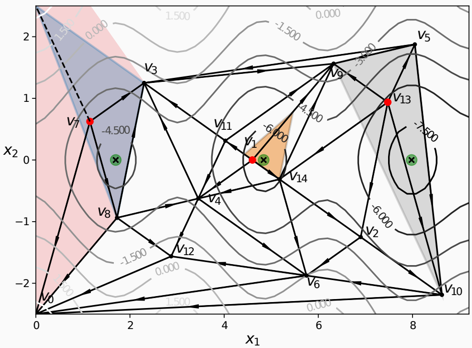
- At this point are two possibilities:
  1. If we have $n+1$ vertices with incidence on an edge $\overline{v_i v_j} \subseteq \mathcal{H}^1$ in every required Cartesian direction then we have a simplex within $\mathrm{st}\,(v_i)$ with a complete Sperner labelling
  2. In the case where we do not have $n+1$ vertices in every required section then by construction there is no vertex between v_i and the boundary of f defined by Ω in the required section. The two possibilities are:
     2.1 In the case where the constraint is not active and there exists at least one point v_k boundary where ∇f does not point towards the boundary and by the MVT v_k can receive a unique Sperner label from which we can construct a simplex within $\mathrm{st}\,(v_i)$ with Sperner labelling
     2.2 In the case where the constraint is active a local minimum lies on the constraint which is in the domain defined $\mathrm{st}\,(v_i)$

Backup slides: Overview of proof of the stationary point theorem v

- Following the combinatorial version of Brouwer's fixed point theorem [Henle, 1979] since ∇f is continuous and the domain $\mathrm{st}\,(v_i)$ is compact we can produce a sequence of complete triangulations with arbitrarily small size in which the size of the simplices decreases toward zero
- This sequence produces a sequence of vertices with gradients $\nabla f(V)$ pointing in every $n+1$ direction. By continuity there is a vector $\nabla f(\mathbf{X})$ near the sequences, since the zero vector is the only vector pointing in all $n+1$ directions we have a point $\mathbf{X}$ bounded by the domain defined by $\mathrm{st}\,(v_i)$ where $\nabla f(\mathbf{X}) = \bar{0}$

This concludes the proof.

Backup slides: Overview of proof of the stationary point theorem vi



Backup slides: Overview of proof of the stationary point theorem vii

- The three circled crosses are the (approximate) minimima of the objective function within the given bounds.
- Here we have divided the plane so that the 3 required directions are $[0, \frac{\pi}{2})$, $[\frac{\pi}{2}, \pi)$ and $[\pi, 2\pi)$
- Note that this division is arbitrary and any $n + 1 = 3$ subdivisions can be chosen as long as all possible $n + 1 = 3$ directions that can form a simplex in the space are covered (affinely independent)
- The three possible Sperner simplices are contained within the star domains of each minimiser $\mathrm{st}\,(v_1)$, $\mathrm{st}\,(v_7)$ and $\mathrm{st}\,(v_{13})$
    1. v_7 is an example of a simplex without a complete Sperner labelling the red shaded area around v_7 is the bounded domain wherein at least one local minimum exist

Backup slides: Overview of proof of the stationary point theorem viii

2. v_{13} has three possible edges in $[\frac{\pi}{2}, \pi)$ on which a point exists that can be used as a vertex to receive a Sperner labelling for that direction namely $\overline{v_{13}v_{14}}$, $\overline{v_{13}v_2}$ and $\overline{v_{13}v_{10}}$. The only possible edges in the $[0, \frac{\pi}{2})$, $[\frac{\pi}{2}, \pi)$ directions are $\overline{v_{13}v_5}$ and $\overline{v_{13}v_9}$ respectively. The simplex $\overline{v_5v_9v_{10}}$ drawn in the figure is not necessarily the simplex with a Sperner labelling. The three vertices of the Sperner simplex which are proven to exist through the MVT exists on each of the edges $\overline{v_{13}v_{14}}$, $\overline{v_{13}v_2}$ and $\overline{v_{13}v_{10}}$ in a subdomain of this simplex $\overline{v_5v_9v_{10}}$
3. v_1 for example the simplex surrounding the minimiser is a possible Sperner simplex with vertices on the edges in every required direction

Backup slides: Overview of proof of the stationary point theorem ix

- Note that if the edge $\overline{v_{13}v_{14}}$ was chosen instead of $\overline{v_{13}v_{10}}$ then the local minimum of the function would be outside the domain of the simplex with the Sperner labelling. This is an important observation because it demonstrates that the theorem cannot be used to further refine the location of the local minimum from the domain $\mathrm{st}\,(v_{13})$ using mechanisms of the proof, it only states that at least one local minimum exists within $\mathrm{st}\,(v_{13})$
- The boundaries of $\mathrm{st}\,(v_{13})$ can be found using the $3-$chain $C_{13}(\mathcal{H}^3)$ of simplices in $\mathrm{st}\,(v_{13})$, recall that the directions of simplices higher than dimension 2 are undefined and so the directions can be arbitrarily chosen

$$C_{13}(\mathcal{H}^3) = \overline{v_{13}v_{10}v_5} + \overline{v_{13}v_5v_9} + \overline{v_{13}v_9v_{14}} + \overline{v_{13}v_{14}v_2} + \overline{v_{13}v_2v_{10}}$$

Backup slides: Overview of proof of the stationary point theorem x

- $C_{13}(\mathcal{H}^3)$ clearly forms a cycle, applying the boundary operator we find the faces defining the bounds of the domain of $\mathrm{st}\,(v_i)$ which in this case is the chain of edges with defined direction

$$\partial(C_{13}(\mathcal{H}^3)) = -\overline{v_{10}v_5} + \overline{v_5v_9} - \overline{v_9v_{14}} + \overline{v_{14}v_2} + \overline{v_2v_{10}}$$

thus $\partial\left(\partial(C(\mathcal{H}^3))\right) = \emptyset$

Backup slides: Overview of proof of the compact invariance theorem i

Theorem

*(**Invariance of an adequately sampled simplicial complex $\mathcal{H}$**) For a given continuous objective function f that is adequately sampled by a sampling set of size N. If the cardinality of the minimiser pool extracted from the directed simplex $\mathcal{H}$ is $|\mathcal{M}|$. Then **any further increase of the sampling set N will not increase $|\mathcal{M}|$**.*

Backup slides: Overview of proof of the compact invariance theorem ii

Definition

Consider a simplicial complex $\mathcal{H}$ built on an objective function f with a compact feasible set Ω using Definitions **??** through 5. The surface is said to be **adequately sampled** if there is one and only one true stationary point within every domain defined by the stationary point theorem

For black box functions there is no way to know if the number and distribution of sampling points is adequate without more information (for example if the number of local minima are known in the problem).

Backup slides: Overview of proof of the compact invariance theorem iii

First we will prove invariance in the case where $\Omega = [\mathbf{l}, \mathbf{u}]^n$ (ie a compact space)

Overview of *proof* :

- The proof relies on a homomorphism between the simplicial complex $\mathcal{H}$ constructed in the bounded hyperrectangle Ω and the homology (mod 2) groups of a constructed surface $\mathcal{S}$ on which we can invoke the invariance theorem
- Define the *n*-torus $\mathcal{S}_0$ from the compact, bounded hyperrectangle Ω by identification of the opposite faces and all extreme vertices
- Now for every strict local minimum point $\mathbf{p} \in \Omega$ puncture a hypersphere and after appropriate identification the resulting *n*-dimensional manifold $\mathcal{S}_g$ is a connected g sum of g tori $\mathcal{S}_g := \mathcal{S}_0 \# \mathcal{S}_1 \# \cdots \# \mathcal{S}_{g-1} \qquad (g \text{ times})$

Backup slides: Overview of proof of the compact invariance theorem iv

- Any triangulation $\mathcal{K}$ of the topological space $\mathcal{S}$ is homeomorphic to $\mathcal{S}$,

$$\mathbf{H}_k(\mathcal{K}) \cong \mathbf{H}_k(\mathcal{S}) \; \forall k \subset \mathbb{Z}$$

- Note that this homomorphism is for a mod 2 homology between a triangulation $\mathcal{K}$ and the surface $\mathcal{S}$ and is thus undirected
- A triangulation corresponding to all vertices (0-simplices) and faces (simplices) of $\mathcal{K}$ can be directed according to the first 3 definitions for h providing the directed simplicial complex $\mathcal{H}$

- By construction we have, for an adequately sampled simplicial complex $\mathcal{H}$, an equality which exists between the cardinality of $\mathcal{M}$ and the Betti numbers of $\mathcal{S}$ as

$$|\mathcal{M}| = h_1 = \text{rank}(\mathbf{H}_1(\mathcal{S})) = \text{rank}(\mathbf{H}_1(\mathcal{K}))$$

Backup slides: Overview of proof of the compact invariance theorem v

- Here we invoke the invariance theorem

Theorem

*(**Invariance theorem [Henle, 1979]**) The homology groups associated with a triangulation $\mathcal{K}$ of the a compact, connected surface $\mathcal{S}$ are independent of $\mathcal{K}$. In other words, the groups $\mathbf{H}_0(\mathcal{K})$, $\mathbf{H}_1(\mathcal{K})$ and $\mathbf{H}_2(\mathcal{K})$ do not depend on the simplices, incidence coefficients, or anything else arising from the choice of the particular triangulation $\mathcal{K}$; they depend only on the surface $\mathcal{S}$ itself.*

- The invariance theorem can be extended to higher dimensional triangulable spaces using singular homology through the Eilenberg-Steenrod Axioms [Eilenberg and Steenrod, 1952, Henle, 1979]

Backup slides: Overview of proof of the compact invariance theorem vi

- As a direct consequence any triangulation of $\mathcal{S}$ will produce the same homology groups for $\mathcal{K}$
- Adding any new sampling point within the corresponding subdomains of $\mathrm{st}\,(v_i)\ \forall i(v_i \in \mathcal{M} \subseteq \mathcal{H}^0)$ as defined in the stationary point theorem will by the first 4 definitions of h need to be connected directly to v_i by a new edge or the triangulation is no longer a simplicial complex and thus not increase $|\mathcal{M}|$ since only one vertex will be the new minimiser
- After adding any sampling point outside a domain $\mathrm{st}\,(v_i)$ then, through the established homomorphism, any construction of $\mathcal{H}$ will produce the same homology groups since $\mathrm{rank}(\mathbf{H}_1(\mathcal{K}))$ remains unchanged and it is thus not possible for a new vertex to be wrongly identified as a minimiser in the triangulation $\mathcal{H}$

Backup slides: Overview of proof of the compact invariance theorem vii

This concludes the proof that any increase in N will not further increase $|\mathcal{M}|$.

N.B.

Any further refinement in the simplicial complex by further sampling does not increase the number of locally convex sub-domains in a compact space!

Backup slides: Overview of proof of the strong invariance theorem i

Finally we prove a stronger invariance and convergence

- Consider the case where the constraints $\mathbf{g}$ are non-linear
- In addition we allow the objective function f to be non-continuous and non-linear
- It is still assumed that the variables $\mathbf{x}$ are bounded
- Furthermore we assume that there is a feasible solution so that $\Omega \neq \emptyset$ and that there exists at least point in range of f mapped within the domain Ω
- We will prove that if the simplicial sampling sequence [Endres, 16] is used, then shgo-simplicial will **retain the Invariance property**
- Secondly **convergence** of the shgo algorithm to the global minimum is proved if the sub-triangulation simplicial sampling sequence is used

Backup slides: Overview of proof of the strong invariance theorem ii

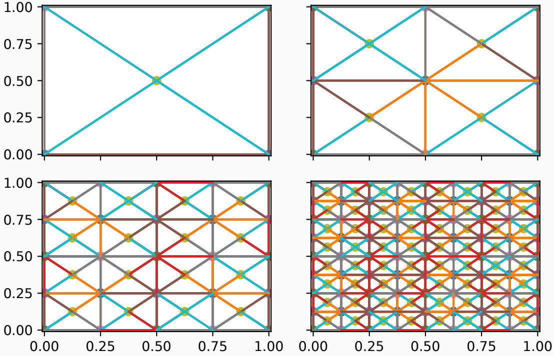


Figure 20: Simplicial sampling by sub-triangulation of hyper-rectangles

Backup slides: Overview of proof of the strong invariance theorem iii

- Before proving these properties we will need to define a new construction to deal with discontinuities in f
- From the definitions of h it is clear that f will only map a subset of the feasible domain Ω, therefore only points within the this domain need to be considered
- A new construction that considers discontinuities (such as singularities) on the hypersurface of f is now defined:

Backup slides: Overview of proof of the strong invariance theorem iv

Definition

For an objective function f, $\mathcal{F}$ is the set of scalar outputs mapped by the objective function $f : \mathcal{P} \rightarrow \mathcal{F}$ for a given sampling set $\mathcal{P} \subseteq \Omega \subseteq \mathbb{R}^n$. If a mapping of a vertex v_i does not exist, then we define the mapping as $f : v_i \rightarrow \infty$. Any such point is excluded from the set $\mathcal{M}$.

Note that any vertex v, $f(v) = \infty$ that is connected to another vertex in Ω that maps to a finite value will never be a minimiser.

Backup slides: Overview of proof of the strong invariance theorem v

Theorem

*(**Invariance of an adequately sampled simplicial complex $\mathcal{H}$ in a non-convex, non-compact space Ω**) For a given non-continuous, non-linear objective function f that is adequately sampled by a sampling set of size N. If the cardinality of the minimiser pool extracted from the directed simplex $\mathcal{H}$ is $|\mathcal{M}|$. Then **any further increase of the sampling set N will not increase $|\mathcal{M}|$**.*

Backup slides: Overview of proof of the strong invariance theorem vi

Overview of *proof* :

- The compact invariance theorem holds for any compact hyperrectangular space $\mathbb{B}_0 = [x_l^1, x_u^1] \times [x_l^2, x_u^2] \times \cdots \times [x_l^n, x_u^n]$
- Consider a set of subspaces $\mathbb{B}_i \cong \mathbb{B}_0$ with $\mathbb{B}_i \subseteq \Omega \ \forall i \in \mathbf{I}$
- That is, $\mathbb{B}_i$ is any compact, rectangular subspace of Ω that is homeomorphic to $\mathbb{B}_0$ (which is also homeomorphic to a point) and can, therefore, be shrunk or expanded to arbitrary sizes while retaining compactness
- Therefore any triangulation $\mathcal{K}_i$ of $\mathbb{B}_i$ retains the compact Invariance property
- We allow all $\mathbb{B}_i$ to be connected or disconnected subspaces with respect to any other $\mathbb{B}_{j \in I}$ within Ω

Backup slides: Overview of proof of the strong invariance theorem vii

- Now consider the (mod 2) homology groups $\mathbf{H}_1(\mathcal{K}_i)$ of $\mathcal{K}_i$
- Since the homology groups are abelian groups the rank is additive over arbitrary direct sums:

$$\operatorname{rank}\left(\bigoplus_{i\in I}\mathbf{H}_1(\mathcal{K}_i)\right) = \sum_{i\in I}\operatorname{rank}(\mathbf{H}_1(\mathcal{K}_i))$$

- Therefore the triangulations of both connected and disconnected subspaces $\mathbb{B}_i$ within a possibly non-compact space Ω will retain the same total rank
- After adequate sampling, the rank of $\mathbf{H}_1(\mathcal{K}_i)$ will not increase by the compact Invariance theorem

Backup slides: Overview of proof of the strong invariance theorem viii

- Any point that is not in Ω is not connected to any graph structure by the definitions in h and therefore cannot increase the rank of any homology group $\mathbf{H}_1(\mathcal{K}_i)$
- Finally any vertex $v_i \in \Omega$ for which $f(v_i)$ does not exist will by the new infinity construction for h be mapped to infinity by the defined mapping $f : v_i \to \infty$
- By the definition, v_i can not be a minimiser and therefore cannot increase the rank of any homology group $\mathbf{H}_1(\mathcal{K}_i)$
- We have shown that the total rank of the homology groups triangulated on all connected and disconnected subspaces $\mathbb{B}_i \in \Omega$ will not increase after adequate sampling
- It remains to be proven that these subspaces exist within Ω

Backup slides: Overview of proof of the strong invariance theorem ix

- We adapt the convergence proof used by [Paulavičius et al., 2014] for subdivided simplicial complexes

Proposition

For any point $\mathbf{x} \in \Omega$ and any $\epsilon > 0$ there exists an iteration $k(\epsilon) \geq 1$ and a point $\mathbf{x}_i^k \in \mathcal{H}^n \in \Omega$ such that $\left\|\mathbf{x}_i^k - \mathbf{x}\right\| < \epsilon$.

- Sampling points $\mathbf{x}_i$ are vertices $\mathcal{H}^0$ belonging to the set of n-dimensional simplices $\mathcal{H}^n$
- Let δ_{max}^k be the largest diameter of the largest simplex
- Since the subdivision is symmetrical all simplices have the same diameter δ_{max}^k after every iteration of the complex
- At every iteration the diameter will be divided through the longest edge, thus reducing the simplices' volumes

Backup slides: Overview of proof of the strong invariance theorem x

- After a sufficiently large number of iterations all simplices will have the diameter smaller than ϵ
- Therefore the vertices of the complex will converge to any and all points inside compact subspaces $\mathbb{B}_i$ within Ω
- Since we have assumed that $\Omega \neq \emptyset$ this proves the existence of subspaces $\mathbb{B}_i$

This concludes the proof.

Convergence

From this proof the **convergence to a global minimum within Ω**, if it exists, also trivially follows by noting that $\mathbb{B}_i$ is homeomorphic to a point and that the stationary point theorem applies to any minimiser in $\mathbb{B}_i$. In practice the definition of h is implemented in [Endres, 16] by using exception handling that can capture any mathematical errors in addition to converting any none float numbers outputted by an objective function to infinity objects.

Backup slides: Overview of proof of the strong invariance theorem xii

Example

We expand the bounds of the Ursem01 function for two dimensions [Gavana, 2016]

$$\min f, \;\; x \in [0, 10] \times [0, 10]$$

Subject to the following non-linear constraints:

$$\begin{aligned}
(x_1 - 5)^2 + (x_2 - 5)^2 + 5\sqrt{x_1 x_2} - 29 &\geq 0 \\
(x_1 - 6)^4 - x_2 + 2 &\geq 0 \\
9 - x_2 &\geq 0
\end{aligned}$$

$$f(\mathbf{x}) = -\sin(2x_1 - 0.5\pi) - 3\cos(x_2) - 0.5x_1$$

Backup slides: Overview of proof of the strong invariance theorem xiii

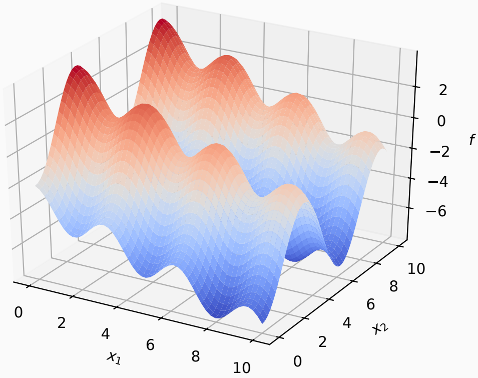
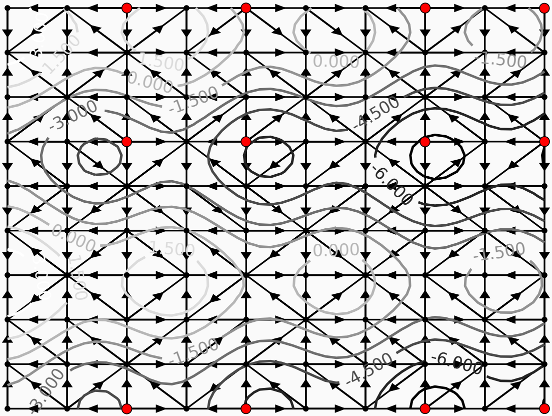


Figure 21: 3-dimensional plot of the Ursem01 function with expanded bounds

Backup slides: Overview of proof of the strong invariance theorem xiv

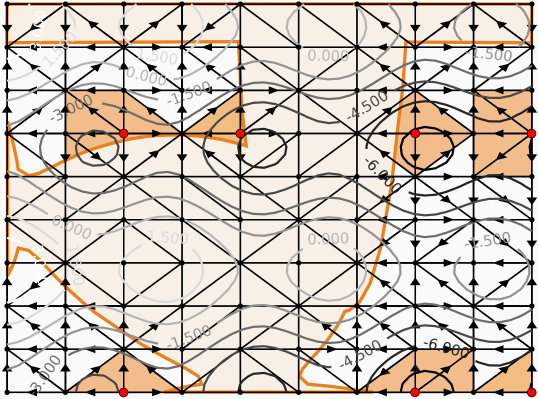
First consider $\mathcal{H}$ without the non-linear bounds, here $|\mathcal{M}| = 12$:



Backup slides: Overview of proof of the strong invariance theorem xv

After applying the non-linear version of h, the non-linear bounds produce the following disconnected simplicial complexes:

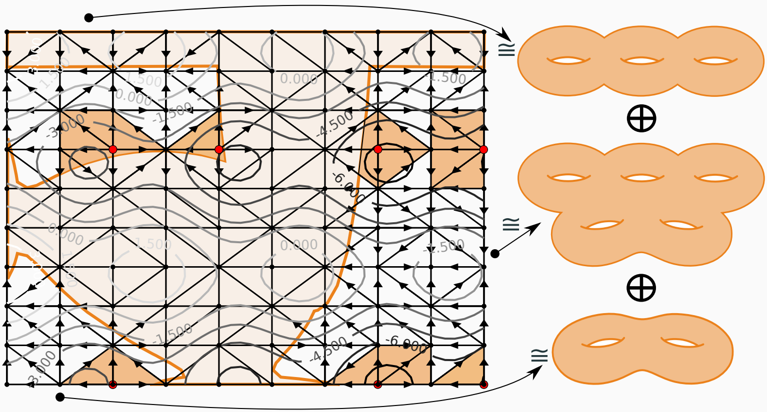
Backup slides: Overview of proof of the strong invariance theorem xvi



Backup slides: Overview of proof of the strong invariance theorem xvii

We use the fact that for abelian homology groups the rank is additive over arbitrary direct sums $\mathbf{rank}\left(\bigoplus_{i\in I} \mathbf{H_1}(\mathcal{K}_i)\right) = \sum_{i\in I} \mathbf{rank}(\mathbf{H_1}(\mathcal{K}_i))$:

Backup slides: Overview of proof of the strong invariance theorem xviii



Backup slides: References to obscure theorems and other additional information sources i

- Discrete MVT: https://www.sciencedirect.com/science/article/pii/S0377221707009952 . https://www.maa.org/sites/default/files/0746834259610.di020780.02p0372v.pdf . https://www.maa.org/sites/default/files/0746834259610.di020780.02p0372v.pdf . https://en.wikipedia.org/wiki/Mean_value_theorem#Mean_value_theorem_in_several_variables (NOTE: The proof provided here is based on Lipschitz continuity)

Backup slides: Backup figures i

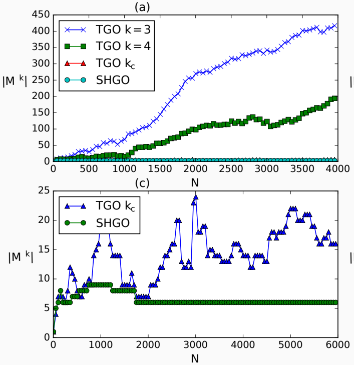


Figure 22: Invariance of homology groups after adequate sampling