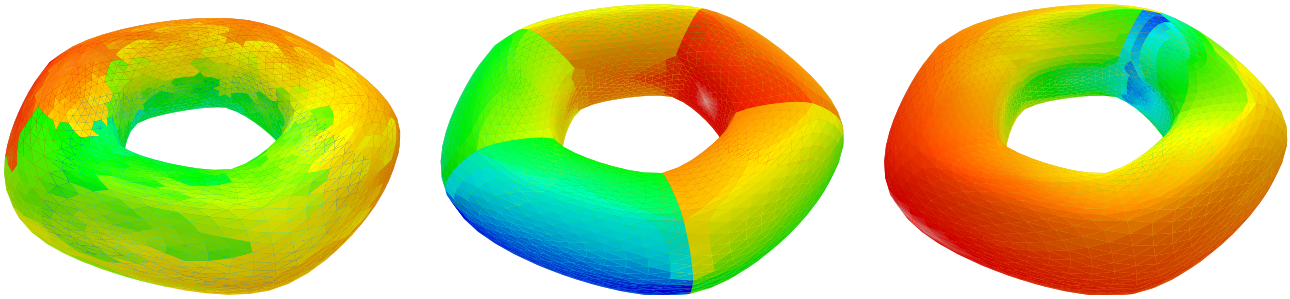
Towards optimality in discrete Morse theory

THOMAS LEWINER, HÉLIO LOPES AND GEOVAN TAVARES

Department of Mathematics — Pontifícia Universidade Católica — Rio de Janeiro — Brazil
{tomlew, lopes, tavares}@mat.puc--rio.br.

Abstract. Morse theory is a fundamental tool for investigating the topology of smooth manifolds. This tool has been extended to discrete structures by Forman, which allows combinatorial analysis and direct computation. This theory relies on discrete gradient vector fields, whose critical elements describe the topology of the structure. The purpose of this work is to construct optimal discrete gradient vector fields, where optimality means having the minimum number of critical elements. The problem is equivalently stated in terms of maximal hyperforests of hypergraphs. Deduced from this theoretical result, a construction algorithm is provided. The optimal parts of the algorithm are proved, and the part of exponential complexity is replaced by heuristics. Although reaching optimality is MAX–SNP hard, the experiments on odd topological models are almost always optimal.
Keywords: ***Morse Theory. Forman Theory. Computational Topology. Discrete Mathematics.***



(a) No constraint (b) Minimal 2–norm (c) Minimal Z coordinate

Figure 1: *Discrete gradient vector fields with geometrical constraints.*

1 Introduction

Morse theory [16] is a fundamental tool for investigating the topology of smooth manifolds. Particularly for computer graphics, many applications have been deduced from the smooth case [8, 12, 13, 18]. Morse proved that the topology of a manifold is very closely related to the critical points of a real smooth map defined on it. The simplest example of this relationship is the fact that if the manifold is compact, then any continuous function defined on it must have a maximum and a minimum. Morse theory provides a significant refinement of this observation.

Forman's discrete Morse theory. The recent insights in Morse theory by Forman [3, 4] extended several aspects of this fundamental tool to discrete structures. Its combinatorial aspect allows computation completely independent of a geometric realization: the algorithms we designed do not require any coordinate or floating–point calculation, and geometrical constraints can be applied independently. Forman proved several results and provided many applications of his theory [5, 6]. Once a Morse function has been defined on a smooth manifold, then information about its topology can be partly deduced from its critical points (i.e. where the gradient vanishes).

Optimality. Similarly to the differential case, Forman proved that the topology of a CW–complex can be partly read out of the critical cells of a discrete gradient vector field defined on it. The topological information will be concise if the discrete gradient vector field has few critical cells. Hence, we will say that a gradient vector field is *optimal* if it has the minimum possible number of critical cells.

Results. The goal of this work is to build optimal gradient vector field. To do so, we develop in section 4 *Structure of a discrete gradient vector field* a hypergraph representation of discrete gradient vector field. We introduce the notion of hyperforest (section 4(c) *Hyperforests*) and prove the equivalence between discrete gradient vector fields and hyperforests in theorem 18. We stated the equivalent of a critical cell for a hyperforest in proposition 20. We finally prove that the

Preprint MAT. 11/02, communicated on August 1^{st}, 2002 to the Department of Mathematics, Pontifícia Universidade Católica — Rio de Janeiro, Brazil. The corresponding work was published in Experimental Mathematics, volume 12, numer 3, pp. 271–285. A.K. Peters, 2003.

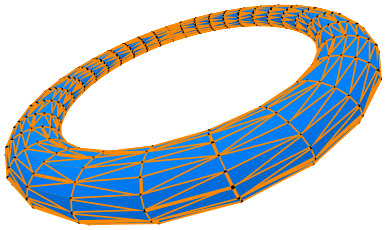


Figure 2: *A triangulated torus.*

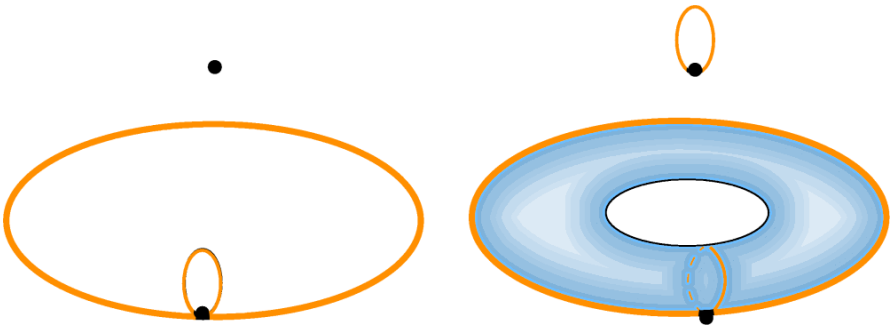


Figure 3: *A construction of a torus with 4 cells.*

minimum number of critical cells is a topological invariant for 3–manifolds in theorem 24.

We provide in section 5 *Constructing discrete gradient vector fields* an algorithm to build discrete gradient vector field on general cell complexes of arbitrary dimension. This algorithm is worst–case quadratic in execution time. It is not guaranteed to be optimal, but it gives optimal results in most of the cases (see section 6(c) *Experiments*). For the particular case of 2–manifolds, the algorithm is proven to be optimal, although the general problem is MAX–SNP hard [10]: a MAX–SNP hard problem is an NP–hard problem for which any polynomial algorithm can lead arbitrary far from the optimum.

Outline. This paper is organized as follow. Section 2 *Discrete Structures* and section 3 *Forman's Discrete Morse Theory* recall some basic definitions of discrete Morse theory. The theoretical base of our work is detailed in section 4 *Structure of a discrete gradient vector field*, and the optimality of our construction is described in section 5 *Constructing discrete gradient vector fields*. As the problem is NP hard, we used the heuristics discussed with their experimental results in section 6 *Further heuristics*.

2 Discrete Structures

(a) Cell Complexes

A cell complex is, roughly speaking, a generalization of the structures used to represent solid models: it is a consistent collection of cells (vertices, edges, faces…). In particular, triangulations of topological spaces or 3D meshes are cell complexes (see Figure 2). Figure 3 gives a minimal construction of a torus cell complex. A complete introduction to cell complexes can be found in [15].

Definition 1 (Cell) *A* cell $\alpha^{(p)}$ *of dimension* p *is a set homeomorphic to the open* p*–ball* $\{x \in \mathbb{R}^p : \|x\| < 1\}$.

When the dimension p of the cell is obvious, we will simply denote α instead of $\alpha^{(p)}$.

Definition 2 (CW–complex) *A* CW–complex K *is built by starting off with a discrete collection of 0–cells (vertices) called* K^0*, then attaching 1–cells (edges) to* K^0 *along their boundaries, obtaining* K^1*, then attaching 2–cells (faces) to* K^1 *along their boundaries, writing* K^2 *for the new space, and so on, giving spaces* K^n *for every* n.

A CW–complex will be said to be finite when it is built out of a finite number of cells. In this work, we will consider only finite (and thus regular) CW–complexes. This permits to compute them.

A p–cell $\alpha^{(p)}$ is a face of a q–cell $\beta^{(q)}$ $(p < q)$ if $\alpha \subset closure\,(\beta)$. If $q = p - 1$, we will use the notation $\alpha^{(p)} \prec \beta^{(q)}$, and say that α and β are *incident*.

In a sense, a cell complex is a generalization of a graph, as a graph can be seen as a cell complex of dimension 1. Nevertheless, we can also represent a cell complex by a multigraph (a multigraph admits multiple links), called the *Hasse diagram*.

Definition 3 (Hasse diagram) *The* Hasse diagram *of a cell complex* K *is the oriented multigraph* H*:*

- *Each node of* H *represents a cell of* K.
- *The links of* H *joins nodes representing incident cells of* K*. The source node of each link is the one of highest dimension.*

The Hasse diagram is usually drawn with the nodes ranked by their dimension. On Figure 4, the faces (2–cells) are aligned on top rank, the edges (1–cells) on the middle one and the vertices (0–cells) on the bottom rank. A link between two nodes symbolizes that the corresponding cells are incident.

(b) Hypergraphs

In the dual graph of a non–manifold cell complex, links that join more than two nodes may appear. This would not fit in the definition of a simple graph, but in the following one:

Definition 4 (Hypergraph) *A* hypergraph *is a pair* (N, L)*, where* L *is a family of families of* N*. The elements of* L *are called* hyperlinks.

We will classify non–empty hyperlinks into the *regular hyperlinks* (or shortly, *link*), which join two distinct nodes as

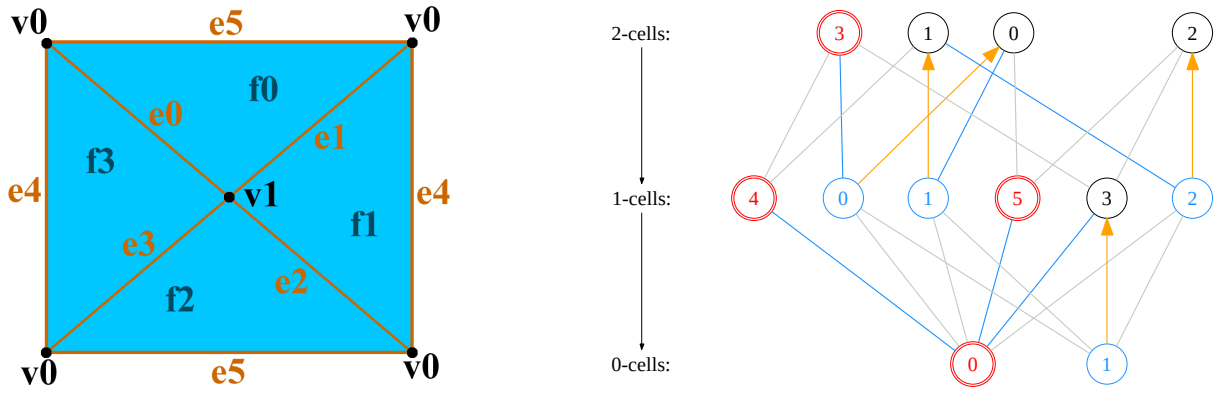


Figure 4: *The Hasse diagram of a non–PL torus.*

in simple graphs, the *loops*, which are incident to only one node, and the *non–regular hyperlinks*, which either join three or more nodes or are multiply incident to one node. We can extract the simple graph part of a hypergraph by considering its *regular components*:

Definition 5 (Regular components) *The* regular components *of a hypergraph* (N, L) *are the connected components of the simple graph* (N, R)*, where* R *is the set of the regular hyperlinks of* (N, L)*.*

We will give a hypergraph a simple orientation by distinguishes one node of each hyperlink as its *source node*.

A hypergraph can be represented by a bipartite graph [1]. This gives a simple but expensive representation of hypergraphs:

Definition 6 (Bipartite graph of a hypergraph) *The* bipartite graph $\mathcal{B}(H)$ *of a hypergraph* H *is the simple graph whose class of nodes are the nodes of* H *on one side, and the links of* H *on the other side.*
For every hyperlink l *of* H*, there are* $\#l$ *links of* $\mathcal{B}(H)$ *joining the node representing* l *in* $\mathcal{B}(H)$ *to each representing node, in* $\mathcal{B}(H)$*, incident to* l *in* H*.*

When H is oriented, $\mathcal{B}(H)$ will be oriented the following way:

- If a node n of H is the source of a hyperlink l, then the node representing l will be the source of the link of $\mathcal{B}(H)$ joining n to l.
- If a node n of H is the not source of an incident hyperlink l, then its representing node in $\mathcal{B}(H)$ is the source node of the link joining it to the representing node of l in $\mathcal{B}(H)$

The operation of taking the bipartite graph of a hypergraph can be reversed. Depending on which class of nodes becomes the links of the hypergraph, we can obtain a hypergraph or its dual. The bipartite graph is not supposed to give a consistent orientation in the general case. Therefore, the hypergraph representing a bipartite graph will not always be oriented.

Definition 7 (Hypergraphs of a bipartite graph) *A bipartite graph* B *admits two representations by hypergraphs:* $\mathcal{B}^{-1}(B)$ *and its dual* $\mathcal{D}\left(\mathcal{B}^{-1}(B)\right)$*. The nodes of* $\mathcal{B}^{-1}(B)$ *represents the nodes of one class of* B*, and the links of* $\mathcal{B}^{-1}(B)$ *represents the nodes of the other class. For every node* l *of the second class, there is a hyperlink of* $\mathcal{B}^{-1}(B)$ *joining all the nodes adjacent to* l*.*

3 Forman's Discrete Morse Theory

(a) Discrete gradient vector field

Definition 8 (Combinatorial vector field) *A* combinatorial vector field $\mathcal{V}$ *defined on a cell complex* K *is a collection of disjoint pairs of incident cells* $\{\alpha^{(p)} \prec \beta^{(p+1)}\}$*.*

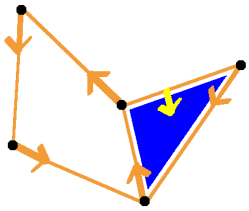


Figure 5: *An example of a combinatorial vector field.*

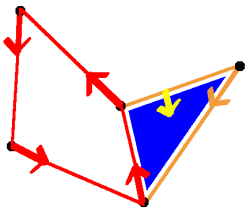


Figure 6: *The closed* $\mathcal{V}$*–path of the vector field of Figure 5 (in red).*

We will represent a combinatorial vector field by an arrow from the cell of lower dimension to its paired cell of higher dimension (see Figure 5).

Definition 9 ($\mathcal{V}$–path) *A* $\mathcal{V}$–path *is an alternating sequence of cells* $\alpha_0^{(p)}$, $\beta_0^{(p+1)}$, …, $\alpha_r^{(p)}$, $\beta_r^{(p+1)}$, $\alpha_{r+1}^{(p)}$ *satisfying:* $\{\alpha_i^{(p)} \prec \beta_i^{(p+1)}\} \in \mathcal{V}$ *and* $\alpha_{i+1}^{(p)} \neq \alpha_i^{(p)}$*.*

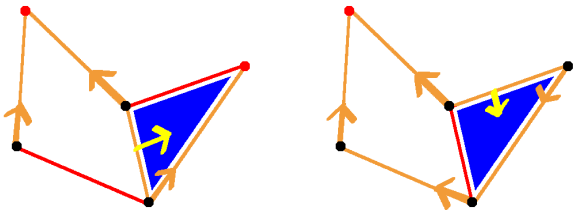
A $\mathcal{V}$–path is non–trivial and closed if $r \geq 1$ and $\alpha_{r+1} = \alpha_0$. For example, Figure 6 shows in red the closed $\mathcal{V}$–path of the combinatorial vector field of Figure 5.

Definition 10 (Discrete gradient vector field) *A* discrete gradient vector field *is a combinatorial vector field with no non–trivial closed $\mathcal{V}$–path.*

(b) Critical cells

Morse proved that the topology of a manifold is related to the critical elements of a smooth function defined on it. Forman gave an analogous result, with the following definition for the critical cells:

Definition 11 (Critical cells) *A cell α is* critical *if it is not paired with any other cell.*



(a) valid (b) optimal

Figure 7: *Examples of discrete gradient vector field.*

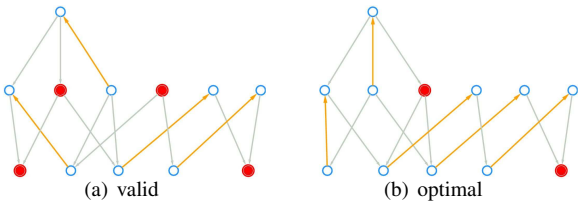
The example of Figure 5 is not a discrete gradient vector field as it contains a closed $\mathcal{V}$–path. On Figure 7, the critical cells of the discrete gradient vector field are drawn in red.

The number of critical cells is not a topological invariant of the cell complex, as it depends on the discrete gradient vector field considered. For example, with an empty discrete vector field (i.e. no cells are paired) every cell is critical, which would be the maximal number of critical cells. In this work, we are more concerned in minimizing this number, as it would give a more concise description of the topology.

(c) Hasse diagram of vector fields

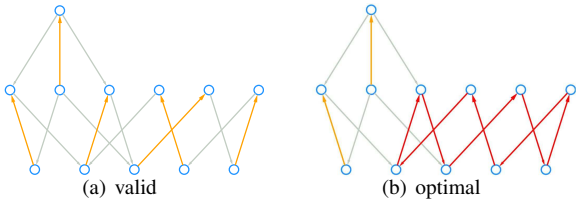
A combinatorial vector field is a matching in the Hasse diagram: each pair of $\mathcal{V}$ corresponds to matched nodes in the Hasse diagram.

We will represent such a matching by inverting the orientation of the link between each pair in $\mathcal{V}$: the arrow's source node will be $\alpha^{(p)}$ for each $\{\alpha^{(p)} \prec \beta^{(p+1)}\} \in \mathcal{V}$. For example, Figure 8 shows the Hasse diagram of the discrete gradient vector fields of Figure 7.



(a) valid (b) optimal

Figure 8: *Hasse diagram of the examples of Figure 7.*



(a) valid (b) optimal

Figure 9: *The Hasse diagram of the combinatorial vector field of Figure 5, and the circuit of its closed $\mathcal{V}$–path (in red).*

With this modified orientation, a closed $\mathcal{V}$–path is just an oriented circuit in the Hasse diagram (see Figure 9). A discrete gradient vector field contains no closed $\mathcal{V}$–path, and thus will be an *acyclic matching*.

4 Structure of a discrete gradient vector field

A discrete gradient vector field has been defined as an acyclic matching in the Hasse diagram (see section 3(c) *Hasse diagram of vector fields*). This involves two problems: creating a matching, and removing cycles. Those two problems are separately well understood (see [14] for matching theory, and [9] for graph algorithms). However, when combined, they create NP hard problems [11]. In this Section, we will give another point of view on discrete Morse theory in terms of the simplest (linear instead of quadratic) of those two problems: creating forests. We will prove our combined problem can be seen as a *hyperforest* creation problem.

(a) Layers of the Hasse diagram

In an n–combinatorial manifold, a (n-1)–cell is incident to either 1 or 2 n–cells [15]. So the dual layer n/(n-1) of the Hasse diagram will be represented by a pseudograph (a pseudograph admits loops and multiple links), called the dual pseudograph. This pseudograph can be seen as the hypergraph representation of the *dual layer* n/(n-1) of the Hasse diagram.

Definition 12 (Layer of the Hasse diagram) *The* layer p/q *of the Hasse diagram, $|p-q|=1$, of a cell complex K is an oriented simple bipartite graph. Its classes of nodes are the p– and q–cells of K. Its links joins nodes representing incident p– and q–cells of K.*

This definition puts a difference between layers of type p/q and q/p. The orientation of those layers is the same as the one of the original Hasse diagram. For example, Figure 11 shows the hypergraph of the layer 2/1 of the Hasse diagram of Figure 10.

(b) Reduced layer of a combinatorial vector field

Considering successive layers of the Hasse diagram is redundant: each p-cell of K appears in general in 4 layers (p/(p+1), p/(p-1), (p+1)/p, (p-1)/p). When the Hasse diagram is oriented by a discrete gradient vector field (see section 3(c) *Hasse diagram of vector fields*), the matching splits on the different layers. The following reduction allows such partition:

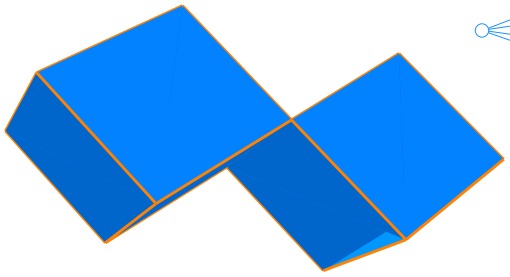


Figure 10: *A double cube.*

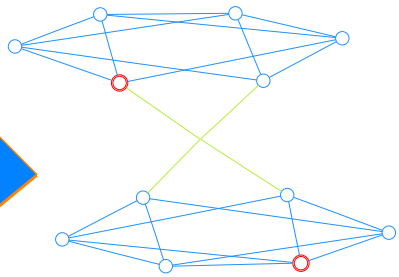


Figure 11: *The 2/1–hypergraph of Figure 10.*

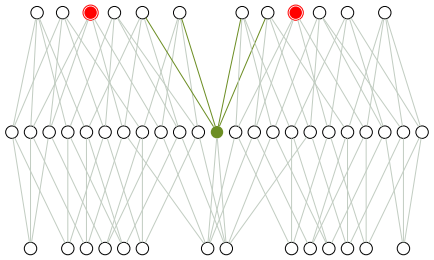


Figure 12: *The Hasse diagram of the double cube of Figure 10.*

Definition 13 (Reduced layers of a combinatorial vector field) *Let K be a cell complex, $\mathcal{V}$ a combinatorial vector field defined on it and B the layer p/q of the Hasse diagram oriented by $\mathcal{V}$, with $|p-q|=1$. The* reduced layer B' *is an oriented bipartite graph obtained by removing from B:*

- *the p–cells of K paired with a q'–cell of K in $\mathcal{V}$, $q' \neq q$.*
- *the q–cells of K unpaired or paired with a p'–cell in $\mathcal{V}$, $p' \neq p$.*

Notice that any $\mathcal{V}$–path is entirely represented in one of the reduced layers. For example, Figure 13 shows in blue the edges in the Hasse diagram of Figure 10 that belongs to the reduced layer 2/1. The corresponding hypergraph (see section 4(d) *Discrete gradient vector field and hyperforests*) is a forest (see Figure 14).

(c) Hyperforests

A forest is a graph with no circuit. Here is a natural extension of forests for hypergraphs:

Definition 14 (Oriented hypercircuit) *An* oriented hypercircuit *in a hypergraph is a sequence of distinct nodes $n_0, n_1, \ldots, n_{r+1}$ such that $n_{r+1} = n_0$ and for all $0 \leq i \leq r$, n_i is the source of a hyperlink leading to n_{i+1}.*

Definition 15 (Hyperforest) *We will say that a simply oriented hypergraph is a* hyperforest *if each node is the source of at most one hyperlink, and if it does not contain any hypercircuit.*

On Figure 15 for example, we can see how the non-regular hyperlink (in green) form a kind of forest. We can deduce from this definition the following properties:

Proposition 16 *Let HF be a hyperforest, and R one of its regular components.*

(i) The regular components of a HF are simple trees.

(ii) There is at most one node in R which is the source of either a loop or non–regular hyperlink.

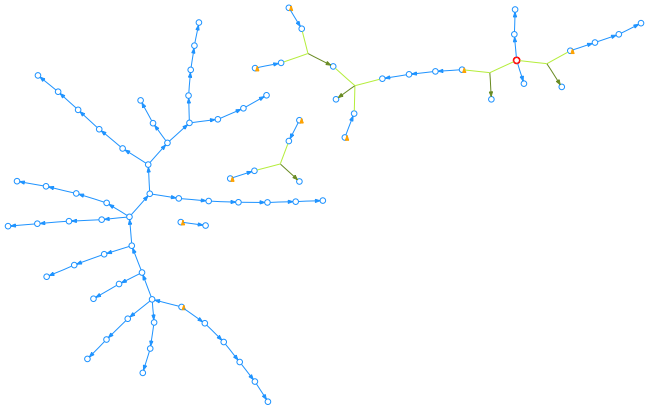


Figure 15: *A part of the hyperforest 2/1 resulting while processing a solid 2-sphere.*

Proof: **(i)**. Suppose R had a (simple) circuit $n_0, n_1, \ldots, n_{r+1} = n_0$. There are $(r+1)$ nodes and $(r+1)$ regular links in this circuit. As a node cannot be the source of two links, each node is the source of exactly one link.
Suppose, without loss of generality, that n_0 is the source of the link $\{n_0, n_1\}$. n_1 is incident to two links of the circuit: $\{n_0, n_1\}$ and $\{n_1, n_2\}$. As it is not the source of the first one, so it is the source of $\{n_1, n_2\}$. Continuing those deductions, we prove that all the links of the circuit are oriented in such a way to form an oriented hypercircuit.
As HF is a hyperforest, this leads to a contradiction. Therefore, R is a simple tree.

(ii). Let k be the number of nodes of R. As R is a tree, it has $(k-1)$ (regular) links. The end nodes of those links are nodes of R, as those links are regular (see definition 5). Therefore, among those k nodes, there are $(k-1)$ nodes that are the source of regular hyperlinks. So there is at most $k-(k-1)=1$ node in R which is the source of either a loop or a non–regular hyperlink. ■

(d) Discrete gradient vector field and hyperforests

We defined a discrete gradient vector field as an acyclic matching in the Hasse diagram (see section 3(c) *Hasse dia-*

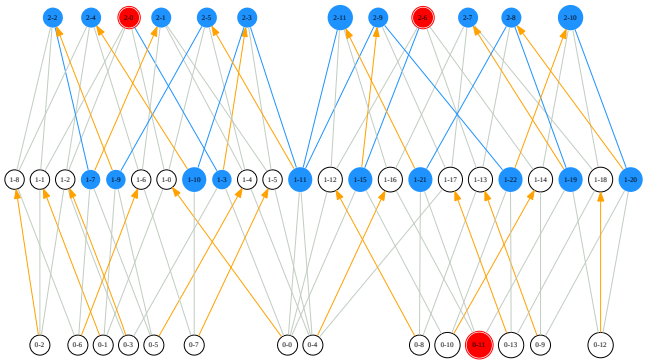


Figure 13: *The reduced layer 2/1 of the double cube (blue nodes).*

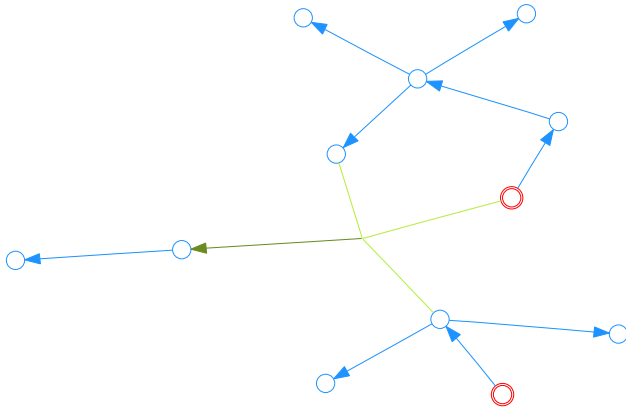


Figure 14: *The hypergraph of the reduced layer 2/1 of Figure 13.*

gram of vector fields), and a hyperforest as a hypergraph without hypercircuit. It seems natural to see a discrete gradient vector field as a collection of hyperforests, extracted from the hypergraphs of the different layers of the Hasse diagram.

Definition 17 (Hypergraphs of a combinatorial vector field) *Let K be a cell complex, $\mathcal{V}$ a combinatorial vector field defined on it and B' the reduced layer p/q of $\mathcal{V}$ ($|p-q|=1$). The p/q–hypergraph of $\mathcal{V}$, H, is the hypergraph representation of B': $H = \mathcal{B}^{-1}(B')$. H is oriented as follow: the source node of a hyperlink of H is the node representing its paired cell in $\mathcal{V}$.*

For example, Figure 16 shows the hyperforest corresponding to a Hasse diagram.

Theorem 18 *Let $\mathcal{V}$ be a combinatorial vector field. $\mathcal{V}$ is a discrete gradient vector field on an n–cell complex K if and only if the 0/1, 1/2, ...(n-1)/n hypergraphs of $\mathcal{V}$ are hyperforests.*

As the dual of a hyperforest is a hyperforest, the theorem is valid for any sequence obtained by replacing the p/q–hypergraph by the q/p–hypergraph of $\mathcal{V}$.

Proof: The orientation of HF ensures the first condition of definition 15. As any $\mathcal{V}$–path is entirely represented in one of the reduced layers, we just need to prove that a closed $\mathcal{V}$–path is a hypercircuit in one of the hypergraphs.

Let $n_0, n_1, \ldots, n_{r+1} = n_0$ be an oriented hypercircuit in the p/q–hypergraph HF. From definition 14, n_i is the source of a hyperlink l_i incident to n_{i+1}. This hyperlink l_i represents a q–cell β_i of K, and n_i represents a p–cell α_i. As n_i is the source of l_i, we know form the orientation of definition 17 that α_i and β_i are incident and form a pair in $\mathcal{V}$. So $\alpha_0, \beta_0, \ldots, \alpha_r, \beta_r, \alpha_{r+1}$ is a $\mathcal{V}$–path. As $n_{r+1} = n_0$ and $r \geq 1$, this is a closed $\mathcal{V}$–path.

This argument can be reversed to prove that a closed $\mathcal{V}$–path is hypercircuit in one of the p/q–hypergraphs. ■

We will now define the analogue of critical cells for hyperforests. This will be the foundation of the algorithm of section 5 *Constructing discrete gradient vector fields*. A critical element of a discrete gradient vector field will be represented by a regular component of one of its hyperforest.

Definition 19 (Critical component) *A regular component of a hyperforest will be called* critical *if none of its nodes is the source of either a loop or a non–regular hyperlink.*

Proposition 20 *Let HF be the p/q–hyperforest of K. The number of critical components of HF is exactly the number of critical p–cells of K.*

Proof: Every possible critical p–cell is represented HF and its corresponding reduced layer B'. The isolated nodes of B' are not matched with any cell of K, and remain isolated nodes in HF. Those nodes are critical components, according to definition 19.

We know from proposition 16 that each regular component R is a simple tree. In such a tree with k nodes, there are $(k-1)$ (regular) links. All links are oriented, so among those k nodes, $(k-1)$ are the sources of links of R, and is therefore not critical. If R is not a critical component, there is exactly one node of R which is the source of either a loop or a non–regular hyperlink, i.e. it is not critical.

If R is a critical component, this node is neither the source of a loop nor of a non–regular hyperlink. From the definition 5 of a regular component, this node is not incident to any regular hyperlink not in R. All those links of R are already paired with other nodes. So this node is unpaired in B'. From definition 13, it cannot be paired with a cell outside B'. Therefore, it is an unpaired node, i.e. a critical cell. ■

(e) Optimality of hyperforests

An optimal discrete gradient vector field will have the minimal possible number of critical components in each hyperforest extracted from the p/q layer. There are as many non–critical elements in a hyperforest as its number of hyperlinks (non–critical elements are paired with an incident hyperlink). Therefore, an optimal discrete gradient vector field

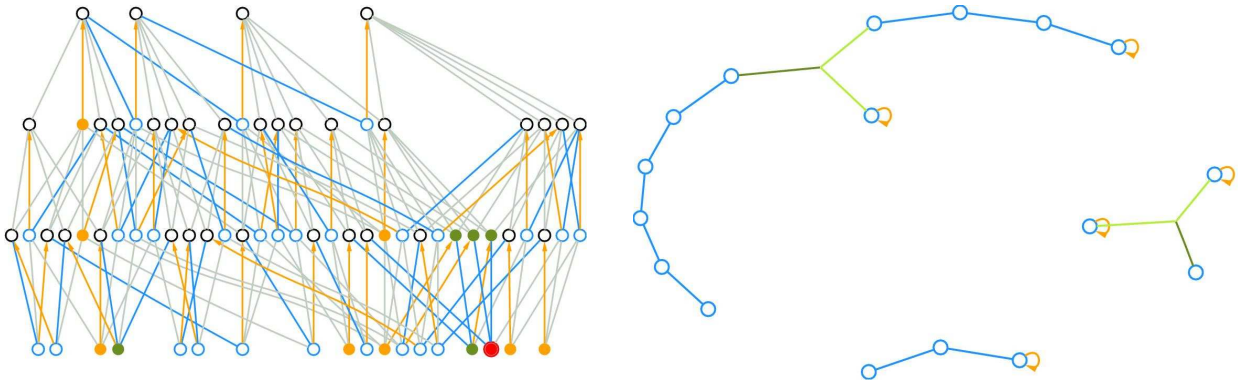


Figure 16: *The Hasse diagram of a discrete gradient vector field on a 4 cubes solid model and the corresponding 1/0–hyperforest.*

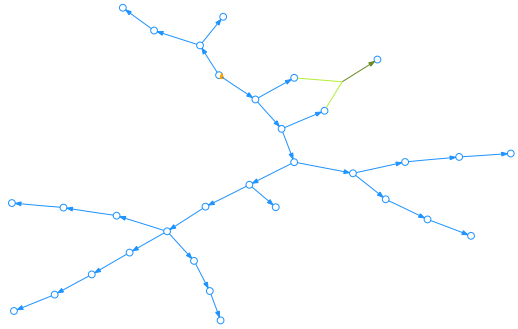


Figure 17: *Detail of a hyperlink insertion in the dual hyperforest appearing with a solid torus model.*

has the maximum number of hyperlinks in each of its hyperforests. This proves, by the way, that the problem of finding a maximal hyperforest in a hypergraph is also MAX–SNP hard (see [10]). For example, adding the hyperlink on the left side of Figure 17 allows us to pair it with the node on the left. Thus, there will be less critical (unpaired) nodes.

(f) A topological invariant for 3–manifolds

Definition 21 (Discrete Morse numbers) *The Morse number $\mathcal{M}_p(K)$ of index p of a cell complex K is the minimum possible number of critical p–cells, considering all possible discrete gradient vector field defined on K.*

Morse theory is linked to simple homotopy. To prove the invariance of the Morse numbers, we could prove that topologically equivalent cell complexes are simple homotopic, and that simple homotopic spaces have the same discrete Morse numbers. Unfortunately, the first affirmation is not true in the general case. However, it states for 3–manifolds. We will use the following theorems, which demonstration can be found respectively in [17] and [2, 25.1].

Theorem 22 (3–Manifold Hauptvermutung) *Any two triangulations of a topological 3–manifold have a common subdivision.*

Theorem 23 *If K_* is a subdivision of K, then K and K_* are simple–homotopy equivalent.*

The proof of the invariance now follows:

Theorem 24 (Invariance of discrete Morse numbers) *Let K and L be homeomorphic 3–manifolds. Then for all p, $\mathcal{M}_p(K) = \mathcal{M}_p(L)$.*

Proof: Let $\mathcal{V}$ be an optimal discrete gradient vector field defined on K. We will prove the theorem by the absurd. Suppose L would have its Morse number of index n higher than the number of critical n–cells of $\mathcal{V}$. We will construct a discrete gradient vector field $\mathcal{W}$ on L with the same number of critical elements as $\mathcal{V}$.

From theorem 22, there exists a common subdivision to K and L. We deduce from theorem 23 that L can be obtained from K by a finite number of collapses and extensions.

If M_* is an extension of M, and $\mathcal{V}$ is a discrete gradient vector field defined on M, we know from [4, section 12] that we can define a discrete gradient vector field $\mathcal{V}_*$ on M_* with the same number of critical elements as $\mathcal{V}$. If M collapses on M_*, we know from [4, lemma 4.3] that we can extend $\mathcal{V}_*$ on M without adding any critical element.

Therefore, we can build a discrete gradient vector field $\mathcal{W}$ on L with the same number of critical elements as $\mathcal{V}$. This leads to the desired contradiction. ■

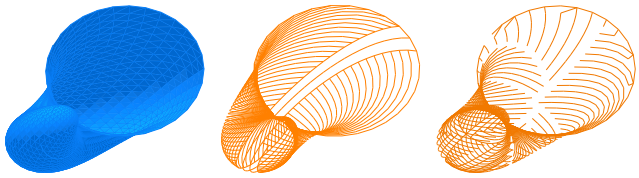
5 Constructing discrete gradient vector fields

The algorithms process each layer or dual layer of the Hasse diagram. For each of those, they define a hyperforest extracted from the layer's hypergraph. Then they give an orientation to those hyperforests, i.e. they define a discrete gradient vector field.

(a) Validity of local optimization

We proved in section 4(f) *A topological invariant for 3–manifolds* that the minimal number of critical cells is an invariant at least for 3–manifolds. Therefore, maximizing the number of hyperlinks in each layer of the Hasse diagram gives a global maximum:

Consider two different n/(n-1)–hyperforests HF_1 and HF_2 giving the same number of critical cells (or critical components). Now call C_1 and C_2 the two cell complexes



(a) Original model (b) C_1 (c) C_2

Figure 18: *A contractible space and the complement cell complexes C_1 and C_2 of two different hyperforests HF_1 and HF_2 defined on it.*

represented by the cells of dimension $\leq n$ and whose (n-1)–cells are not in HF_1 and HF_2 respectively (see Figure 18). From theorems 3.3 and 3.4 of [4], C_1 and C_2 are simple homotopic and they have the same minimal number of critical cells. We conclude by induction that, in the case of 3–manifolds, maximizing the number of hyperlinks in each hyperforest generates an optimal discrete gradient vector field.

(b) Maximum number of regular hyperlink

Each regular component R of H is determined before any construction of HF. For any hyperforest HF, consider RT the simple graph which nodes are the n nodes of R and which links are the regular hyperlinks of HF incident to those nodes. As R is a regular component of H, there is no regular hyperlink incident to a node of R and a node out of R, so RT is well defined. As HF is a hyperforest, there is no circuit in RT: RT is a collection of k trees. So RT has $(n-k)$ links. The maximum number of links will thus be for k minimal, i.e. RT being a unique (connected) tree. This optimum can be reached by constructing a spanning tree on each regular component of H [9].

(c) Maximizing the number of loops

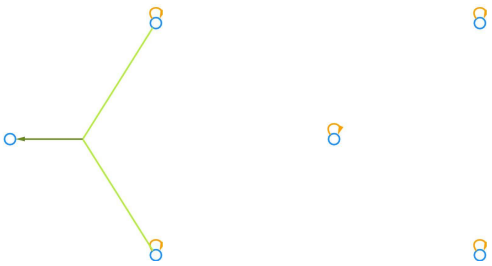


Figure 19: *Replacing a non–regular hyperlink by a loop.*

Each connected component of HF is either critical or incident to either a loop or a non–regular hyperlink. The problem of regular hyperlinks has been resolved optimally, and we want now to maximize the number of loops and non–regular hyperlinks of HF. If a critical component is incident to a loop in H, then adding this loop to HF generates another hyperforest with one more hyperlink and one less critical component. If a regular component is incident to a loop l in H and to a non–regular hyperlink nl in H and HF, then replacing nl by l in HF generates another hyperforest with the same number of hyperlinks (and less risk to create a hypercircuit). This process is illustrated on Figure 19. Therefore, we can always generate a hyperforest HF with the maximum possible number of hyperlinks such that every regular component incident to a loop in H is incident to a loop in HF.

(d) Condition for non–regular hyperlink insertion

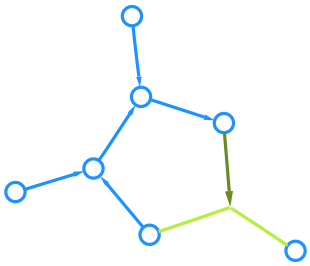


Figure 20: *A hyperlink creating a hypercircuit.*

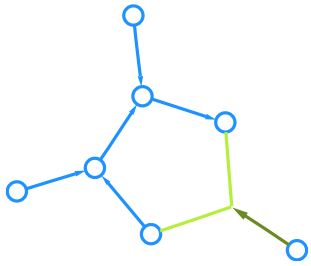


Figure 21: *A regular component incident to only one hyperlink.*

Let HF be the hyperforest being created out of the hypergraph H. There is one critical cell of HF in each of its critical component (proposition 20). A hyperlink can be added to HF only if it is incident to at least one critical component. For a non–regular hyperlink nl, let $\mathcal{C}(nl)$ denote the set of connected components of H containing a critical regular component of H incident to nl in HF.

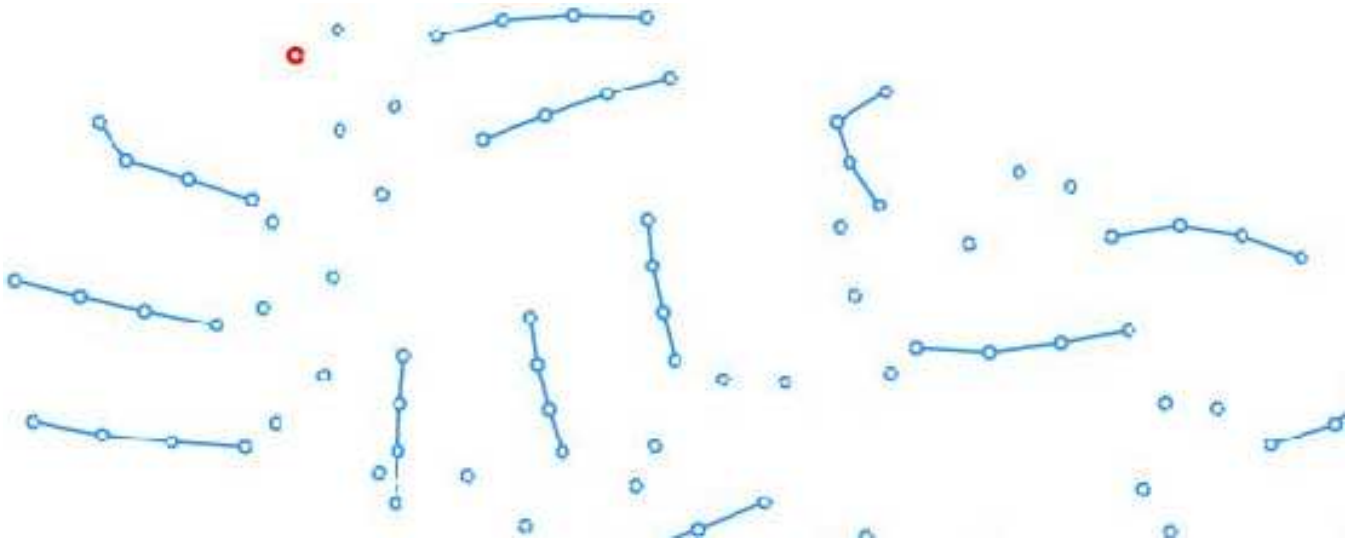
The hyperlink nl can create a hypercircuit in a connected component C of $\mathcal{C}(nl)$ when it is incident to more than one node of C, or more than one time to a node of C, and when the source of nl is a node of C (see Figure 20).

If there exists a connected component C in $\mathcal{C}(nl)$ such that nl is incident to only one node of C, and simply incident, then nl can be added to HF, and one of those nodes will be its source node. If there does not exist such a node, we can remove nl from H as it will never be part of HF. This case is valid if nl is not incident to any critical component, i.e. if $\mathcal{C}(nl)$ is empty.

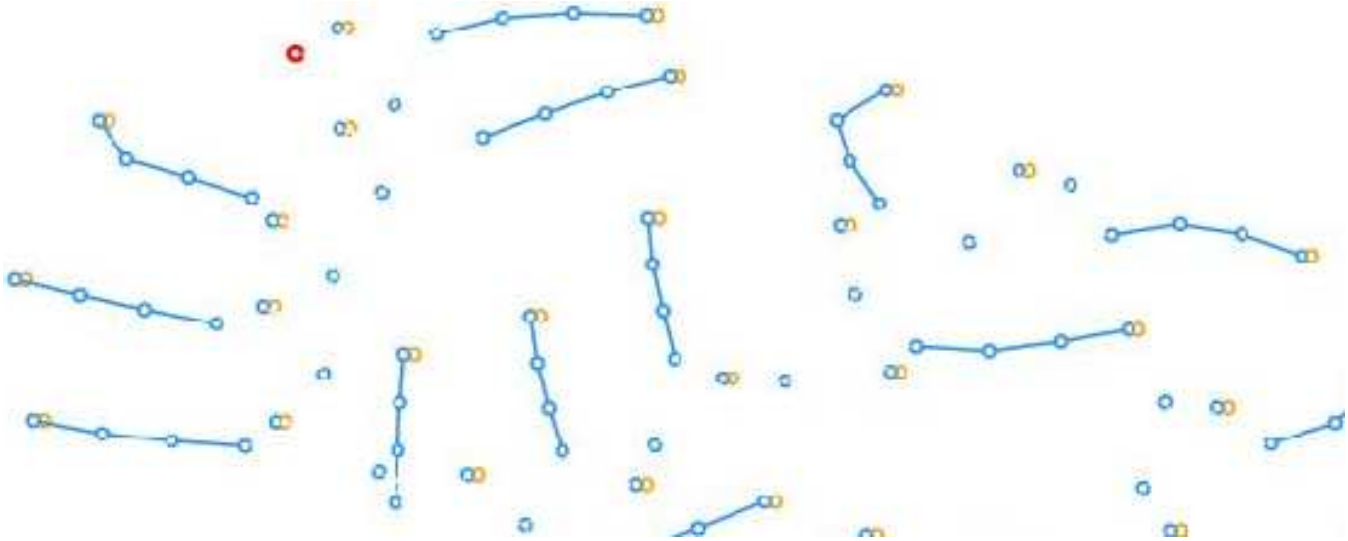
In particular, when a regular component is incident to only one hyperlink and if the hyperlink is incident only once to this regular component we can add it to HF (see Figure 21).

(e) Algorithm outline

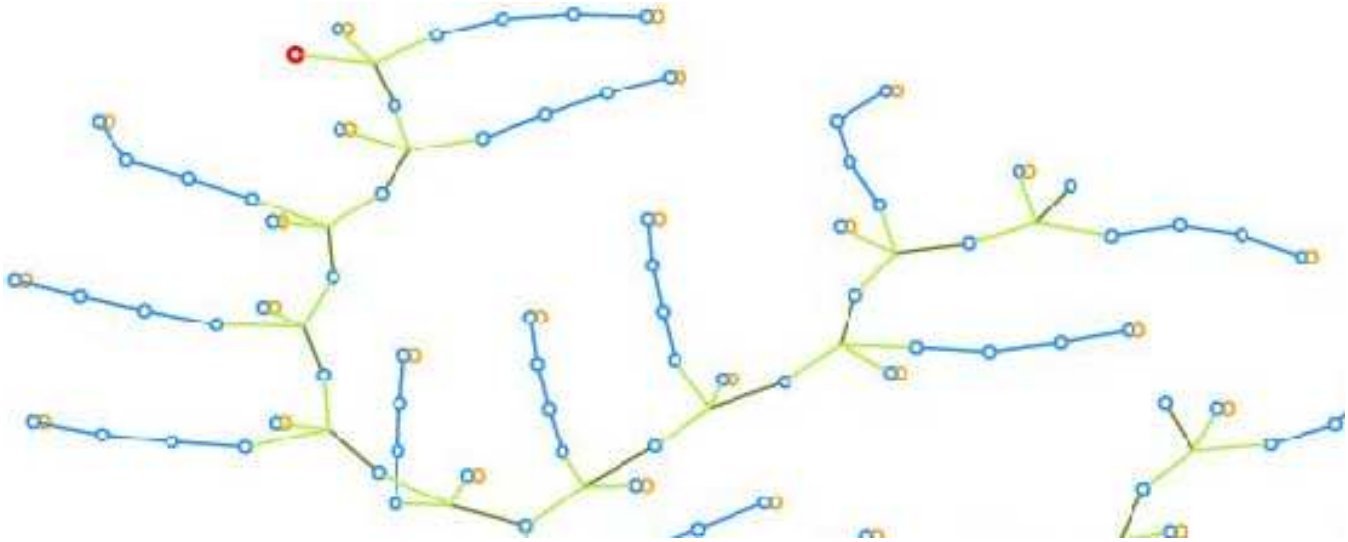
We must first choose which layers of the Hasse diagram we process. In fact, we can process all of them, indifferently from their direct or their dual hypergraph representation. We know that the dual pseudograph of a manifold has no non–regular hyperlink, and that the direct hypergraph of the first layer is a simple graph. Those two simple cases could be useful as the construction of hyperforest is linear on pseudograph and quadratic on general hypergraphs. For



(a) steps 1–3: Forest of the regular components.



(b) step 4: Adding loops.



(c) step 5: Adding non–regular hyperlinks.

Figure 22: *Steps of the algorithm on a part of the hyperforest 2/1 of* $S^2 \times S^1$.

example, a solid model could be processed by the following sequence of layers: 0/1,1/2,3/2; or 3/2,2/1,0/1.

In this work, all the algorithms to extract a hyperforest HF out of a hypergraph H process by the following steps (see Figure 22(a), Figure 22(b) and Figure 22(c)):

1. Initiate HF with the nodes of H.
2. Generate a spanning tree on every regular component of H.
3. Add all the links of those spanning trees to HF.
4. If a regular component is incident to some loops, add one of them to HF.
5. Process the non–regular hyperlinks of H.

The 4 first steps of the algorithm are linear, and guaranteed to be optimal in any case. The last step requires some heuristics as detailed below.

6 Further heuristics

The complexity of the 4 first steps of our algorithm is linear in time of execution. Each of those steps gives an optimal result. However, the problem of finding an optimal discrete gradient vector field is MAX–SNP hard (see [10]). Thus, the last step of our algorithm, i.e. deciding which non–regular hyperlinks of H will be added to HF, must be a much more expensive. If the size of the hypergraph allows it, we could use an exponential algorithm, generating all possible hyperforest and testing which is the maximal one. For the general case, we provide in this Section different heuristics and their results to complete the last step of our algorithm.

(a) Greedy methods

Let HF be the hyperforest being created out of the hypergraph H. We can try to add the hyperlinks of H to HF in order they appear after a sort. The criterion for a hyperlink to be added or not to HF has been discussed in section 5(d) *Condition for non–regular hyperlink insertion*.

The priority on links (which appears in the sort) can be quite arbitrary, as there is no polynomial approximation. We tested 3 of them:

- minimal number of incident regular components.
- minimal number of incident critical components in HF.
- maximal number of incident non–critical components in HF.

The problem that appears with those criterions is that the priority must be calculated again each time a hyperlink is added to HF (as some components change status from critical to non–critical). So the complexity of such a heuristic is quadratic.

(b) Mixing with geometry

As for the algorithm of [11], we can impose some more conditions on our discrete gradient vector fields. However, there is a difference with that case: the geometry can influence the result, as the hyperforest of a layer will be different if the hyperforest of the precedent layer processed is a geometrical minimum.

There are different constraints we can add on our hyperforest HF:

- The spanning tree of the regular components of HF can be chosen to be a minimal spanning tree.
- The loops added to the regular components of HF can minimize the same function, in order to have the root of the spanning trees at a minimal position.
- The roots of the spanning trees of the critical components of HF can also be at a minimal position.
- The priority used in the greedy heuristics (see section 6(a) *Greedy methods*) can be derived from the same geometrical function.

The discrete gradient vector fields generated on a torus model with those constraints are represented on Figure 1 for different geometrical function. The vector field goes from cold colors to hot ones. All of them have 4 critical points (1 vertex, 2 edges, 1 face).

(c) Experiments

	HG Simpl	Min Deg	Min Def	Max Cpl
Direct	1208	530	14	402
Dual	7258	728	8	934
Sym Direct	3580	658	50	702
Sym Dual	3842	566	6	722

Table 1: *Number of redundant critical cells per method on the models of tables 2.*

We compared our different heuristics on two kinds of models: Hachimori's examples [7] (mainly non–constructible, and other solid models at the Mat&Mídia Laboratory (see Table 2). The results of those processes are given on Table 1. The different heuristics we implemented are:

- Direct: processing the layers 0/1, 1/2, 2/3.
- Dual: processing the layers 3/2, 2/1, 1/0.
- Sym Direct: processing the layers 0/1, 1/2, 3/2.
- Sym Dual: processing the layers 3/2, 2/1, 0/1.
- HG Simpl: only simplifying the hypergraph, with no further process.
- Min Def: priority to the hyperlinks incident to the minimum number of critical components.
- Min Deg: priority to the hyperlinks of minimum degree.

The corresponding work was published in Experimental Mathematics, volume 12, numer 3, pp. 271–285. A.K. Peters, 2003.

– Max Cpl: priority to the hyperlinks incident to the maximum number of non–critical components.

Forcing the first and last layers to be processed as 0/1 and n/(n-1), as in the cases of Sym Direct and Sym Dual leads to the best results (see table 2), because it generates fewer non–regular hyperlinks to be processed: the 0/1 layer is a multigraph (1–skeletton) and, for the case of manifolds, the n/(n-1) layer is a pseudograph.

The Sym Dual is usually the best processing order, as it avoids disconnecting the cell complex as would do Sym Direct (for example, a 0/1 spanning tree on a surface with boundary, with a two vertices on the same boundary). In particular for 2–manifolds, the Sym Dual algorithms are optimal [11].

The Min Def priority in the greedy algorithm leads to the best algorithm, from far. When looking at the detailed results, mixing with the geometry of the cell complex, when available, significatively improves the performances of the algorithm.

Figure 23 and Figure 24 show the resulting discrete gradient vector field on a 4D model of a Cartesian product of a Möbius strip by a circle. The algorithm used was the Sym Dual / Min Def algorithm, minmizing the total 'y' coordinate of the spanning trees and root positions. The discrete gradient vector field goes from cold colors to hot ones.

7 Open problems

This work was focused on Forman's discrete Morse theory. We analyzed the building blocs of this theory, and proved the layered structure of discrete gradient vector field. We represented this layer structure by a collection of hyperforests and gave a complete characterization of the critical cells in terms of regular components of hyperforests. We used this analysis to introduce a scheme for constructing discrete gradient vector field on finite cell complexes of arbitrary dimension. Although the general problem is MAX–SNP hard, this construction is quadratic in time in the worst cases, and is proven to be linear and optimal in the case of 2-manifolds. The experimental results showed our algorithm gave an optimal result in most of the cases.

We know from the disproof of the Hauptvermutung (see [17]) that combinatorial invariants of triangulations are not always topological ones. Thus the discrete Morse numbers could not be a topological invariant in the general case. However, for the case of 3–manifolds, we proved here that discrete Morse numbers are topological invariants.

Our algorithms seem to be optimal in all the cases we studied, except for the knotted ball and the Bing's house. The conditions that would guarantee an optimal result in polynomial time remain an open problem.

References

[1] C. Berge. *Graphes et hypergraphes*. Dunod, Paris, 1970.

[2] M. M. Cohen. *A course in simple homotopy theory*. *Graduate text in Mathematics*. Springer, New York, 1973.

[3] R. Forman. A discrete Morse theory for cell complexes. In S. T. Yau, editor, *Geometry, Topology and Physics for Raoul Bott*, pages 112–115. International Press, Cambridge, 1995.

[4] R. Forman. Morse theory for cell complexes. *Advances in Mathematics*, 134:90–145, 1998.

[5] R. Forman. Some applications of combinatorial differential topology. In *Sullivan Fest*, 2001.

[6] R. Forman. A user guide to discrete Morse theory. In *Séminaire Lotharingien de Combinatoire*, volume 48, 2001.

[7] M. Hachimori. Simplicial complex library. infos-hako.sk.tsukuba.ac.jp/~hachi.

[8] J. C. Hart. Morse theory for implicit surface modeling. In H.-C. Hege and K. Polthier, editors, *Visualization and Mathematics*, pages 257–268, Heidelberg, 1998. Springer.

[9] J. Hopcroft and R. E. Tarjan. Efficient algorithms for graph manipulation. *Communications of the ACM*, 16:372–378, 1973.

[10] T. Lewiner. Constructing discrete Morse functions. Master's thesis, *Department of Mathematics, PUC–Rio*, Aug. 2002. Advised by Hélio Lopes and Geovan Tavares.

[11] T. Lewiner, H. Lopes and G. Tavares. Optimal discrete Morse functions for 2–manifolds. *Computational Geometry: Theory and Applications*, 26(3):221–233, 2003.

[12] H. Lopes. Algorithm to build and unbuild 2 and 3 dimensional manifolds. PhD thesis, *Department of Mathematics, PUC–Rio*, 1996. Advised by Geovan Tavares.

[13] H. Lopes and G. Tavares. Structural operators for modeling 3–manifolds. In C. Hoffman and W. Bronsvort, editors, *Solid Modeling and Applications*, pages 10–18. ACM, 1997.

[14] L. Lovász and M. D. Plummer. *Matching theory*. Van Nostrand Reinhold, Amsterdam, 1986.

[15] A. T. Lundell and S. Weingram. *The topology of CW–complexes*. Van Nostrand Reinhold, New York, 1969.

[16] J. W. Milnor. *Morse theory*. Number 51 in *Annals of Mathematics Study*. Princeton University Press, 1963.

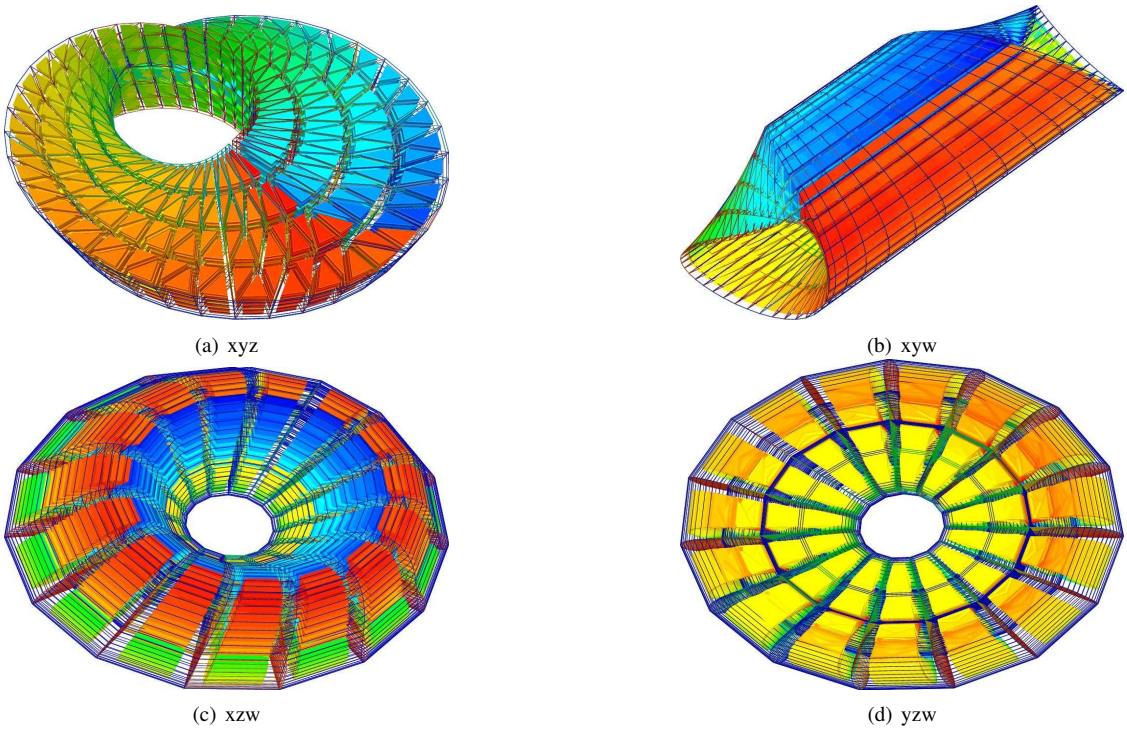
[17] E. E. Moïse. Affine structures in 3–manifolds. *Annals of Mathematics*, 56(2):96–114, 1952.

[18] Y. Shinagawa, T. L. Kunii and Y. L. Kergosien. Surface coding based on Morse theory. *Computer Graphics and Applications*, 11:66–78, 1991.

Preprint MAT. 11/02, communicated on August 1^{st}, 2002 to the Department of Mathematics, Pontifícia Universidade Católica — Rio de Janeiro, Brazil.

Model	Topology	Number of cells	χ	Minimum
bing	3–ball	(480,2511,3586,1554)	1	(1,1,1,0)*
bjorner	projective plane + one facet	(6, 15, 11)	2	(1,0,1)
c-ns	contractible	(12, 37, 26)	1	(1,1,1)
c-ns2	contractible	(13, 39, 27)	1	(1,0,0)
c-ns3	contractible	(10, 31, 22)	1	(1,1,1)
dunce hat	Dunce hat	(8, 24, 17)	1	(1,1,1)
gruenbaum	3–ball	(14, 54, 70, 29)	1	(1,0,0,0)
knot	3–ball	(380,1929,2722,1172)	1	(1,1,1,0)
lockeberg	3–sphere	(12, 60, 96, 48)	0	(1,0,0,1)
mani walkup C	3–sphere	(20, 126, 212, 106)	0	(1,0,0,1)
mani walkup D	3–sphere	(16, 106, 180, 90)	0	(1,0,0,1)
nonextend	contractible	(7, 19, 13)	1	(1,0,0)
poincare	homology sphere	(16, 106, 180, 90)	0	(1,2,2,1)
projective	projective plane	(6, 15, 10)	1	(1,1,1)
rudin	3–ball	(14, 66, 94, 41)	1	(1,0,0,0)
simon	contractible	(7, 20, 14)	1	(1,0,0)
ziegler	3–sphere	(10, 38, 50, 21)	1	(1,0,0,0)
Pile of Cubes	contractible	(572,1477,1266,360)	1	(1,0,0,0)
s2xs1	$S^2 \times S^1$	(192, 588, 612,216)	0	(1,1,1,1)
s3	3–sphere	(162, 522, 576,216)	0	(1,0,0,1)*
solid 2sphere	2–sphere	(64, 144, 108, 26)	2	(1,0,1,0)
Furch knotted ball	3–ball	(600,1580,1350,369)	1	(1,1,1,0)*

Table 2: *Results on solid models. (* points models for which the best result is not obtained by Min Def / Sym Dual but by Min Def / Sym Direct).*

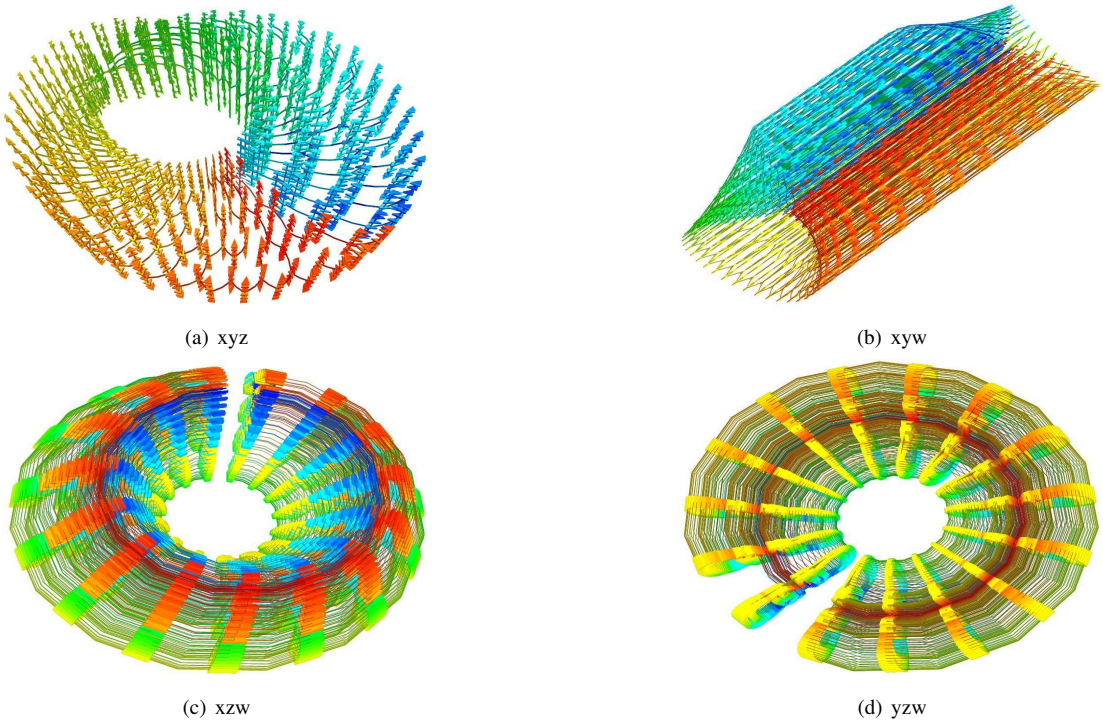


(a) xyz (b) xyw

(c) xzw (d) yzw

Figure 23: *A discrete gradient vector field on a Cartesian product of a Möbius strip by a circle: 2925 0-cells, 10500 1-cells, 12225 2-cells and 4650 3-cells; 1 critical 0-cell, 2 critical 1-cells, 1 critical 2-cell and 0 critical 3-cell. The discrete gradient vector field, drawn for the 1– and 3–cells, goes from cold colors to hot ones. The 3–cells are shrunken for the sake of clarity.*

The corresponding work was published in Experimental Mathematics, volume 12, numer 3, pp. 271–285. A.K. Peters, 2003.



(a) xyz (b) xyw

(c) xzw (d) yzw

Figure 24: *The 3/2–spanning tree of the discrete gradient vector field of Figure 23. The 2–cells of the 3/2–spanning tree are drawn as line, and the 3–cells as small solids.*