Breaking Good: Fracture Modes for Realtime Destruction

SILVIA SELLÁN, University of Toronto
JACK LUONG, California State University, Fresno and University of California, Los Angeles
LETTIJA MATTOS DA SILVA, University of California, Los Angeles and Massachusetts Institute of Technology
ARAVIND RAMAKRISHNAN, University of Maryland and University of Toronto
YUCHUAN YANG, University of California, Los Angeles
ALEC JACOBSON, University of Toronto and Adobe Research

Vibration modes = a shape’s natural deformation patterns

Our Fracture modes = a shape’s natural breaking patterns

Fig. 1. Drawing an analogy with the well-studied vibration modes, we define a shape’s fracture modes, which we can precompute for realtime applications.

Drawing a direct analogy with the well-studied vibration or elastic modes, we introduce an object’s fracture modes, which constitute its preferred or most natural ways of breaking. We formulate a sparsified eigenvalue problem, which we solve iteratively to obtain the \( n \) lowest-energy modes. These can be precomputed for a given shape to obtain a prefracture pattern that can substitute the state of the art for realtime applications at no runtime cost but significantly greater realism. Furthermore, any realtime impact can be projected onto our modes to obtain impact-dependent fracture patterns without the need for any online crack propagation simulation. We not only introduce this theoretically novel concept, but also show its fundamental and practical advantages in a diverse set of examples and contexts.

CCS Concepts:
• Computing methodologies → Physical simulation.

Additional Key Words and Phrases: fracture simulation, modal analysis.

ACM Reference Format:

1 INTRODUCTION

The patterns and fragmentations formed by an object undergoing brittle fracture add richness and realism to destructive simulations. Unfortunately, existing methods for producing the most high-quality realistic fractures (e.g., for the film industry) require hefty simulations too expensive for many realtime applications. An attractive and popular alternative is to rely on precomputed fragmentation patterns at the modeling stage that can be swapped in at run-time when an impact is detected. Existing prefracture methods use geometric heuristics that can produce unrealistic patterns oblivious of an object’s elastic response profile or structural weaknesses (see Figs. 2, 3 and 6). Geometric patterns alone also do not answer which fragments should break-off for a given impact at run-time, inviting difficult to tune heuristics or complete fracture regardless of impact. As a result, these procedural methods find use when fractures are in the background or obscured by particle effects; elsewhere, video game studios may rely on artist-authored fragmentation patterns.

In this paper, we present a method for prefracturing stiff brittle materials which draws a direct analogy to a solid shape’s elastic vibration modes. We compute a shape’s fracture modes\(^1\), which algebraically span the shape’s natural ways of breaking apart. By introducing a continuity objective under a sparsity-inducing norm to the classic vibration modes optimization problem, we identify unique and orthogonal modes of fracture in increasing order of a generalized notion of frequency.

\(^{1}\)Not to be confused with the "three modes of fracture" [Irwin 1957].
The first $k$ fracture modes can be intersected against each other to define a prefracture pattern as a drop-in replacement to existing procedural methods (see Fig. 2). Furthermore, impacts determined at runtime can be efficiently projected onto the linear space of precomputed fracture modes to obtain impact-dependent fracture without the need for costly stress computation or crack propagation.

We demonstrate the theoretical and practical advantages and limitations of our algorithm over existing procedural methods and evaluate its accuracy by qualitatively comparing to existing works in worst-case structural analysis. We showcase the benefits of our algorithm within an off-the-shelf rigid body simulator to produce animations on a diverse set of shapes and impacts. We show the realtime potential of fracture modes with a prototypical interactive 2D application (see Fig. 4 and accompanying video).

2 RELATED WORK
Fracture simulation has an extensive body of previous work; Muguer-cia et al. [2014] provide a thorough survey. Stiff brittle fracture is characterized when little or no perceptible deformation occurs before fracture (i.e., when the object is otherwise rigid). Modeling the dynamic or quasistatic growth of brittle fracture patterns in a high performant way requires not just high spatial resolution but also high temporal resolution at the microsecond scale [Kirugulige et al. 2007], modeling stress concentration and subsequent release. This process has been approximated, for example, using mass-springs [Hirota et al. 1998; Norton et al. 1991], finite elements [Kaufmann et al. 2009; Koschier et al. 2015; O’Brien and Hodgins 1999; Pfaff et al. 2014; Wick et al. 2010], boundary elements [Hahn and Wojtan 2015, 2016; Zhu et al. 2015], and the material-point method [Fan et al. 2022; Wolper et al. 2020, 2019]. While any such method can eventually meet realtime demands by lowering the discretization fidelity (e.g., on low-res cage geometry [Müller et al. 2004; Parker and O’Brien 2009] or a modal subspace [Gloudu et al. 2012]) or assuming large enough computational resources, we instead focus our attention to previous methods which achieve realtime performance via the offline, removing the need for heuristics or data-based approaches.

Creating prefracture patterns manually requires skill and time, precluding fully automatic pipelines. Many commercial packages (e.g., UNREAL ENGINE, Houdini) implement or suggest geometric prefracturing heuristics to segment a shape into solid subfragments. Voronoi diagrams of randomly scattered points [Oh et al. 2012; Raghavachary 2002] capture the stochastic quality of fracture, but result in overly regular and convex fragments with perfectly flat sides. Despite lacking realism, convexity can be advantageous for simulations. For example, it enables realtime collision detection and even offline at massive scales [Zafar et al. 2010]. Beyond collision detection, approximate convex decomposition has been employed as a prefracturing technique, allowing fracture patterns to be applied locally at low cost [Müller et al. 2013]. Müller et al. [2013] rely on manual intervention at multiple stages, making it not comparable as an automatic method; however, we compare to the Voronoi decomposition it is based on in Figs. 2 and 3.

Schvartzman and Otaduy [2014] increase the space of possible fractures beyond convexity by computing the Voronoi decomposition on a higher dimensional embedding. The regularity of these fragments can be augmented further by randomly generated “cutter” objects (i.e., bumpy planes slicing through the input shape) or perturbed level set functions (see, e.g., [Museth et al. 2021]). These stochastic methods often miss obvious structural weaknesses (see Fig. 3) or result in implausible fragments (see Fig. 6). While the defects of these methods can be resolved manually, hidden behind destruction dust or obscured by fast explosions, our fracture modes consistently produce non-convex fractures whose boundaries originate from minimal stress displacements of the shape. Beyond taking into account the physical elastic behavior of the geometry, our method can incorporate constraints to avoid fractures in certain areas. While our fracture modes are slower to compute than geometric-only procedural prefracture, this cost is added only at the offline precomputaton stage.

Given a precomputed fracture pattern, one must decide which fractures are activated when the object receives a given online impact. Strategies range from simple heuristics like Euclidean distance thresholds and centering a spatial fracture pattern on the contact point [Müller et al. 2013; Su et al. 2009] to learning from examples [Schvartzman and Otaduy 2014]. A key practical contribution of our fracture modes is that they span a linear subspace onto which impacts can be cheaply projected to trigger fragment displacements at runtime, removing the need for heuristics or data-based approaches.
Our approach for defining fracture modes lies within a broader class of optimization problems of the form:

$$\arg\min_{X^T MX = 1} \frac{1}{2} \text{trace}(X^T LX) + \sum_{i=1}^{k} g(X_i),$$

(1)

where $X_i$ as the $i^{th}$ column of $X \in \mathbb{R}^{n \times k}$ is referred to as a modal vector or mode, $M$ and $L$ are positive semi-definite $n \times n$ matrices, and $g : \mathbb{R}^n \to \mathbb{R}$ is a sparsity-inducing norm, like $g(x) = \|x\|_1$. When $g \equiv 0$, this reduces to the generalized eigenvalue problem (see e.g., [Bai et al. 2000] Chaps. 4-5), whose solution satisfies $LX = MXA$, where the diagonal $k \times k$ matrix $A$ contains the $k$ smallest eigenvalues ($A_{ii} = X_i^T L X_i$). For non-trivial $g$, we may continue to consider $\lambda_i = X_i^T L X_i + g(X_i)$ as describing the frequency of the $i^{th}$ mode.

Ozolins et al. [2013] proposed the notion of compressed modes using a sparsity-inducing $\ell_1$-norm to compute localized (sparse) solutions to Schrödinger’s equation. Neumann et al. [2014] extended this idea to compressed eigenfunctions of the Laplace-Beltrami operator on 3D surfaces, advocating for an alternating direction method of multipliers (ADMM) optimization method. While ADMM’s standard convergence guarantees [Boyd 2010] do not apply to non-convex problem such as Eq. (1), Neumann et al. [2014] demonstrate successful local convergence albeit with dependency on the initial guess and optimization path. Replacing the $X^T LX$ term with a data-term, Neumann et al. [2013] use a similar ADMM approach to create sparse PCA bases for mesh deformations.

Brandt and Hildebrandt [2017] further extend this line of smooth, sparse modal decompositions by considering $L$ to be the Hessian of an elastic energy. They propose an iterative mode-by-mode optimization. The current mode is optimized by sub-iterations consisting of a quadratic program solve resulting from linearizing the constraints around the current iterate interleaved with normalization in order to approach a unit-norm vector. Despite the conspicuous downside that any sub-optimality of earlier modes is locked in possibly affecting the accuracy of later modes, this method enjoys performance and robustness improvements over the ADMM approach of Neumann et al. [2014]. Therefore, we follow suit with a similarly mode-by-mode fixed-point iteration approach. Unique to our method is that we do not consider the sparsity of the modal vector itself, but rather the sparsity of the mode’s continuity over the domain.

3 FRACTURE MODES

Given an elastic solid object $\Omega \subset \mathbb{R}^d$ and a deformation map $u : \Omega \to \mathbb{R}^d$, we can formulate the object’s total strain energy as

$$E_F(u) = \int_{\Omega} \Psi(u, x) dx,$$

(2)

where $\Psi$ is the strain energy density function evaluated at points $x \in \Omega$ in the undeformed object.

Suppose we allow the deformation map $u$ to fracture the object $\Omega$ into two disjoint pieces $\Omega_1$ and $\Omega_2$ along a given $(d - 1)$-dimensional fracture fault $S$ (see inset). Effectively, we’re allowing $u$ to be discontinuous at $S$. Consider $x_1$ and $x_2$ to be the undeformed positions of points infinitesimally on either side of a point $x \in S$ on the fracture fault. Then, the difference between $u(x_1)$ and $u(x_2)$ describes the pointwise vector-valued discontinuity at $x \in S$:

$$D(u, x \in S) = u(x_1) - u(x_2) \in \mathbb{R}^d,$$

(3)

where $D = 0$ would indicate continuity or absence of fracture. We can then compute the discontinuity energy associated with $S$ as

$$\int_{x \in S} \|D(u, x)\|^2 dx,$$

(4)

which can be seen as a cohesive surface energy from the FEM crack propagation literature (e.g., [Ortiz and Pandolfi 1999]). Unlike similar cohesive energies found in graphics (e.g., for UV mapping [Poranne et al. 2017] or shape interpolation [Zhu et al. 2017]), assuming small displacements affords us this simpler, first-order approximation.

We now consider that the set of admissible discontinuities is not just a single fracture fault, but a finite number of fault patches: $S = \{S_1, \ldots, S_p\}$. We assume that $S$ comes from, for now, arbitrary over-segmentation of $\Omega$. This could be created with a high-resolution Voronoi diagram, by intersecting $\Omega$ with random surfaces, voxel boundaries, or by some a priori distribution of granular subobjects.
We will derive a discrete formulation of the variational problem in \( \ell_\text{Wakin} \) 2008 total discontinuity energy \( E_D(u) = \|D(u, S)\|_{L^1, \Omega} := \sum_{j=1}^{p} \sqrt{\int_{S_j} \|D(u, x)\|^2} dx \). (5)

We add this to the strain energy to form the total energy:
\[
E(u) = E_{\Psi}(u) + \omega E_D(u),
\]
where \( \omega \in \mathbb{R} \) is a positive weight balancing the two terms. In the cohesive FEM context, \( \omega \) can be understood as the square root of the traction-displacement coefficient in the first fracture phase [Chowdhury and Narasimhan 2000].

Minimizing the \( \ell_1 \) norm on the matrix \( D \) is tantamount to minimizing the sparsity-inducing [Candes and Wakin 2008] \( \ell_1 \) norm on the lengths of each row. Minimizing the \( \ell_1 \) norm \( (\sum_j |D_{ij}|) \) directly would also lead to sparsity, but the solution would be rotationally dependent.

We can now define the \( k \) lowest energy fracture modes as the set of mass-orthonormal deformation maps \( \{u_j^k\}_{j=1}^k \) that minimize their combined total energy; i.e.,
\[
\{u_j^k\}_{j=1}^k = \arg\min_{\{u_j\}_{j=1}^k} \sum_{j=1}^k E(u_j), \quad \text{s.t.} \quad \int (u_j)^T \rho u_j dx = \delta^{ij},
\]
where \( \rho \) is the local mass density and \( \delta^{ij} \) is the Kronecker delta. In general, for large enough \( \omega \), minimizers \( u_j^k \) will have exactly zero \( E_D \) on all but a sparse subset of fault patches in \( S \), agreeing with the usual sparse coding theory [Candes and Wakin 2008].

### 3.1 Fracture Modes on Meshes

We will derive a discrete formulation of the variational problem in Eq. (7) for a 2D solid object, represented as a triangle mesh \( \Omega \) with \( n \) vertices and \( m \) faces. In this construction, the mesh’s \( \rho \) interior edges will correspond to admissible fracture faults \( S_1, \ldots, S_p \). Everything that follows is straightforward to extend to 3D solids by exchanging triangles and interior edges for tetrahedra and interior faces.

A traditional piecewise-linear finite element method (FEM) would discretize the strain energy \( E_{\Psi} \) using hat functions \( \phi_j : \Omega \to \mathbb{R} \), \( \forall i = 1, \ldots, n \) and associate a scalar function \( u : \Omega \to \mathbb{R} \) with a vector \( u \in \mathbb{R}^n \) such that
\[
u(x) = \sum_{i=1}^{n} u_i \phi_i(x).
\]
Vector-valued \( u : \Omega \to \mathbb{R}^d \) such as a deformation would be described coordinate-wise in the same way, via a vector \( u \in \mathbb{R}^{dn} \).

Hat functions are by construction continuous. Normally, this is a good thing, but we would like to have functions with arbitrarily large co-dimension one patches of discontinuities. Let us introduce the concept of an exploded mesh \( \hat{\Omega} \), with the same \( m \) faces as \( \Omega \) and same carrying geometry, but where each vertex is effectively repeated for each incident triangle. Thus, \( \hat{\Omega} \) is composed of \( m \) combinatorially disconnected triangles and \( 3m \) vertices \( \tilde{v}_{cf} \) with \( c = 1, 2, 3 \) and \( f = 1, \ldots, m \).

Hat functions \( \tilde{\phi} : \hat{\Omega} \to \mathbb{R} \) defined on \( \hat{\Omega} \) reduce to barycentric coordinate functions extended with zero value outside the corresponding triangle (see inset). These trivially span the space of piecewise linear scalar discontinuous functions \( u \) via vectors \( u \in \mathbb{R}^{3m} \):
\[
u(x) = \sum_{f=1}^{m} \sum_{c=1}^{3} u_{cf} \tilde{\phi}_{cf}(x).
\]

We will use this basis for each coordinate of our vector-valued deformation map, captured in a vector of coefficients \( u \in \mathbb{R}^{(d+1)md} \), with the displacement at vertex \( c \) of face \( f \) selected by \( u_{cf} \in \mathbb{R}^d \).

We may now discretize both terms in our energy in Eq. (6). First, the integral strain energy follows the usual FEM discretization as a sum over each element.
\[
E_{\Psi}(u) = \frac{1}{2} \sum_{f=1}^{m} \int_{f} \Psi(u, x) dx.
\]

We can abstract the choice of \( \Psi \) for now by considering small displacements around the rest configuration such that it can be approximated by its Hessian matrix \( Q \in \mathbb{R}^{(d+1)md \times (d+1)md} \):
\[
E_{\Psi}(u) \approx \frac{1}{2} u^T Qu.
\]

In Section 3.7 we discuss a further approximation specific to our model. Our basis functions \( \tilde{\phi} \) will only allow fractures along the mesh’s interior edges, therefore, the integral in Eq. (5) breaks into a contribution from each interior edge \( e \):
\[
E_e(u) = \sqrt{\int_e \|D(u, x)\|^2} dx,
\]
which we can compute exactly by two-point Gaussian quadrature (the integrand is a second-order polynomial).
For a given edge \( e \) with length \( l \), corresponding to vertex pairs \( \{af, bf\} \) and \( \{cg, dg\} \) (see inset), this amounts to

\[
E_e(u) = \sqrt{\frac{1}{2} \left( \|d(\frac{u_{af}}{\sqrt{v_{af}}} + \frac{d}{\sqrt{v_{bf}}})\|_2^2 + \|d(\frac{u_{cg}}{\sqrt{v_{cg}}} + \frac{d}{\sqrt{v_{dg}}})\|_2^2 \right)}
\]

where

\[
d(t) = \frac{1 + t}{2} (u_{af} - u_{cg}) + \frac{1 - t}{2} (u_{bf} - u_{dg})
\]

measures the pointwise discontinuity for the quadrature at parametric location \( t \in [-1, 1] \) along the edge.

The full discontinuity energy associated with the map \( u(x) \) is given by summing over every interior edge

\[
E_D(u) = \sum_{e=1}^{p} E_e(u).
\]

Finally, we can define the \( k \) lowest-energy discrete fracture modes as column vectors of a matrix \( U \in \mathbb{R}^{(d+1)mdxk} \) satisfying

\[
\arg\min_{U \in \mathbb{R}^{(d+1)mdxk}} \frac{1}{2} \text{trace} (U^TQU) + \varnothing \sum_{l=1}^{k} (E_D(U_l)),
\]

where \( Q \) is the possibly lumped FEM mass matrix defined on the exploded mesh \( \Omega \).

### 3.2 Optimization

The definition of fracture modes on meshes involves solving the optimization problem in Eq. (15). While the objective term is convex, the orthogonality constraints are not. To proceed, we adapt the Iterated Convexification for Compressed Modes (ICCM) scheme proposed by Brandt and Hildebrandt [2017]. ICCM computes the modes sequentially, assuming the first \( i - 1 \) columns of \( U \) have been computed. In the original ICCM formulation, the process for finding the \( i \)th column, \( U_i \) proceeds by choosing a random unit-norm vector \( c \), then repeatedly solving

\[
U_i \leftarrow \arg\min_{U} \frac{1}{2} U^TQU + \varnothing E_D(U)
\]

subject to

\[
\begin{bmatrix}
U_i^T \\
\vdots \\
U_{i-1}^T \\
c^T
\end{bmatrix} \begin{bmatrix}
0 \\
\vdots \\
0 \\
1
\end{bmatrix} = \begin{bmatrix}
0 \\
\vdots \\
0 \\
1
\end{bmatrix}
\]

and updating

\[
c \leftarrow \frac{U_i}{U_i^T MU_i}.
\]

until convergence is detected by \( \|U_i - c\| \) falling below some tolerance \( \varepsilon \). By linearizing the (quadratic) norm constraint, the minimization problem Eq. (16) becomes a convex conic problem that we solve with off-the-shelf techniques (see Appendix A).

We found that random initializations for \( c \) only not introduce non-determinism, but can also sometimes result in a large number of inner iterations and sub-optimal local minima (see Fig. 8). Instead, when computing \( U_i \) we initialize \( c \) with the \( i \)th continuous eigenvectors of \( Q \) (defined on the unexploded mesh). We compute these \( k \) initial vectors at once using the SciPy wrapper for the sparse eigen solver ARPACK [Lehoucq et al. 1998]. We outline our complete fracture mode computation algorithm in Algorithm 1.

### 3.3 Impact-dependent fracture

By construction, the columns of \( U \) form an orthonormal basis of the lowest-energy \( k \)-dimensional subspace of possible fractures of \( \Omega \). This key feature means we can precompute an object’s fracture modes right after its design. Then, inside an interactive application we can project any detected impact onto our modes to obtain impact-dependent realtime fracture (see Fig. 9).

If a collision is detected between \( \Omega \) and another object, with contact point \( p \) and normal \( \hat{n} \), we can define the exploded-vertex-wise impact vector \( w \in \mathbb{R}^{(d+1)md} \). Ideally, \( w \) would be the displacements determined by an extremely short-time-duration simulation of elastic shock propagation. In lieu of being able to compute this in realtime, we use an approximation based on distance to smear the impact into the object:

\[
w_{ef} = g(p, \nu_{ef}) \hat{n}, \quad \forall c = 1, 2, 3, f = 1, \ldots, m
\]

where \( g \) is a filter that vanishes as \( \nu_{ef} \) is far from \( p \). Then, we project \( w \) onto our modes to obtain our projected impact

\[
w^* \equiv \sum_{i=1}^{k} U_i U_i^T Mw = \sum_{i=1}^{k} U_i U_i^T M \delta \hat{n}.
\]

An immediate choice of \( g \) would be a Gaussian density function centered at \( p \). A more physically based choice of \( g \) can be obtained through a single implicit timestep of an elastic shockwave equation

\[
g = \left( M - \tau L \right)^{-1} \hat{M} \delta \hat{n}.
\]
where $\tau$ is the timestep of the simulation our fractures are embedded in. This choice of $g$ has benefits beyond physical inspiration, by ensuring an impact is only blurred onto regions that are geodesically close to one another, regardless of whether they are Euclideanly close (see Fig. 9). However, computing this $g$ upon impact would involve solving a linear system at runtime. We avoid this by precomputing

$$A_i = U_i^T M \left( M - \tau \mathbf{L} \right)^{-1} M,$$

thus requiring only a matrix multiplication at runtime:

$$w^* = \sum_{i=1}^{k} U_i A_i \delta_p \tilde{n}.$$  \hspace{1cm} (23)

Let $\tilde{\Omega}w^*$ be the exploded mesh $\tilde{\Omega}$ as deformed by the map $w^*$. For any two vertices $v_{af}, v_{cg}$ that are coincident in $\tilde{\Omega}$ (i.e., that came from the same original vertex in $\Omega$), we will glue (i.e., un-explode) them if their deformation maps differ by less than some tolerance, $\|w^*_{af} - w^*_{cg}\| < \sigma$. This results in a new fractured mesh $\Omega^*$, whose fracture pattern depends meaningfully on the nature of the impact and which we can output to the simulation.

Our single timestep in Eq. (21) is an approximation that makes $w^*$ depend linearly on the impact. It has the added effect that scaling $\sigma$ and scaling the magnitude of the impact are equivalent in our model. Thus, $\sigma$ could be linked to the force of the impact or the relative speed if one has access to this dynamic information (see Fig. 10). We note that this equivalence is a product of our modeling choices and may not always yield physically accurate results. For example, a large force on a small area may cause immediate local fractures, quickly reducing the stress before it propogates further; in our model, the same impact would likely cause large global fractures.

3.4 Efficient implementation for real-time fracture

In 3D, our fracture mode computation needs a tetrahedralization of the input’s interior, but practical realtime applications prefer to work with triangle surface meshes for input and output. Fortunately, the sparsity inducing discontinuity norm results in fracture modes which are continuous across most pairs of neighboring tetrahedra. It is unnecessary to keep the entire tetrahedral mesh at runtime. Instead, we can determine the connected components determined by neighboring tetrahedra whose shared face’s discontinuity term is below $\sigma$ across all $k$ modes (or below the lowest possible $\sigma$ allowed by the dynamic system). The boundary of each component is a solid

\begin{algorithm}[H]
\caption{Fracture Modes via Adapted ICCM}
\begin{algorithmic}
\State $Q$ be a PSD matrix, $k \in \mathbb{N}$
\State $C \leftarrow$ eigenvectors($Q, M, k$)
\For{$i = 1, \ldots, k$}
\State $c \leftarrow C_i$
\Repeat
\State $U_i \leftarrow$ solve Eq. (16) (see Appendix A)
\State $c \leftarrow U_i / \sqrt{U_i^T M U_i}$
\Until{$\|U_i - c\| \leq \tau$}
\State return $U$
\end{algorithmic}
\end{algorithm}

\begin{algorithm}[H]
\caption{Simple nested cages via binary search}
\begin{algorithmic}
\State $V_{in,F_{in}}$ be the vertex and face lists of the input mesh, and $m_{target}$ the desired number of output faces.
\State \textbf{Binary Search} on offset amount $d$
\State $V_{mc,F_{mc}} \leftarrow$ marching-cubes( distance to $V_{in,F_{in}}$ minus $d$)
\State $V_{d,F_d} \leftarrow$ decimate $V_{mc,F_{mc}}$ to $m$ faces
\State $V_u,F_u \leftarrow$ self-union $V_d,F_d$ via [Zhou et al. 2016]
\State $V_u,F_u \leftarrow$ tetrahedralize $V_u,F_u$ via [Si 2015]
\If {any step failed or $V_i,F_i$ does not strictly contain $V_{in,F_{in}}$}
\State increase $d$
\Else
\State decrease $d$
\EndIf
\State return $V_i,T_i$
\end{algorithmic}
\end{algorithm}

Fig. 11. We use the upper envelope extraction algorithm by Abdrashitov et al. [2021] to obtain smoothed fracture faults.

[Zhou et al. 2016] triangle mesh of a fracture fragment. Since the impact projection described above is linear, we can pre-restrict the projection to vertices on the boundary of these fragments, discarding all internal vertices and the tetrahedral connectivity.

3.5 Simple Nested Cages

In practice, the input model may be very high-resolution, not yet fully modeled when fractures are precomputed, or too messy to easily tetrahedralize. Like many simulation methods before us, we can avoid these potential performance, workflow and robustness problems by working with a tetrahedral coarse cage nesting the input. The \textsc{Nested Cages} method of Sacht et al. [2015] produces tight fitting cages, but suffers long runtimes, potential failure, and may result in a surface mesh which causes subsequent tetrahedralization (e.g., using \textsc{TetWild} [Hu et al. 2018] or \textsc{TetGen} [Si 2015]) to fail.

Therefore, in Algorithm 2 we introduce a very simple caging method inspired by the level-set method of Ben-Chen et al. [2009]. As an example, \textsc{Nested Cages} crashes after a minute on the input mesh in the inset, while our simple algorithm produces a satisfying output cage after 85 seconds.

Like \textsc{Nested Cages}, the output cage will strictly contain the input, but also by construction we ensure that this cage can be successfully tetrahedralized (not just in theory). In a sense, this method provides a different point on the Pareto frontier of tightness-vs-utility. Each step is a fairly standard geometry processing subroutine with predictable performance, and one may even consider using it as an initialization strategy for \textsc{Nested Cages} to improve tightness in the future. We run a max of 10 search iterations, lasting between 5 and 20 seconds each in our examples.
Fracture modes and solid fragment components on the cage’s tetrahedralization can be transferred to the true input geometry by intersecting each connected component against the input mesh. In this way, the exterior surface of each fragment component is exactly a subset of the input mesh.

3.6 Smoothing internal surfaces

By our construction, the fracture boundaries will follow faces of the tetrahedral mesh used for their computation. This reveals aliasing with frequency proportional to the mesh resolution. We may optionally alleviate this by treating each extracted per-tet component membership as a one-hot vector field, which we immediately average onto (unexploded) mesh vertices stored as a matrix $Z \in \{0,1\}^{n \times \text{components}}$, so that $Z_{ij} \in \{0,1\}$ is viewed as the likelihood that vertex $i$ belongs to component $j$. We apply implicit Laplacian smoothing with a time step of $\lambda$ to columns of $Z$:

$$Z \leftarrow (M + \lambda L)^{-1}(MZ),$$  \hspace{1cm} (24)

where $M$, $L$ are the mass and Laplacian matrices, respectively. The resulting $Z$ continue to contain fractional values in $[0,1]$ corresponding to a smoothed likelihood. We now re-extract piecewise-linear (triangle mesh) component boundaries by computing the upper-envelope (tracking the argmax) using the implementation of Abdashitov et al. [2021]. While essentially still using the same tetrahedral mesh, utilizing smoothing and piecewise-linear interpolation greatly reduces aliasing artifacts (see Fig. 11).

By the nature of the Laplacian, Eq. (24) will push our fracture faults towards smooth surfaces. This is in alignment with our modeling decisions at the beginning of Section 3: as our set of possible fault patches $S_i$ becomes larger, the area integral in Eq. (5) will encourage smoother fracture fault surfaces. Thus, the postprocessing described here is not a departure from our model; rather, a way of alleviating the error introduced by the mesh discretization.

In the real world, crystalline materials do break along smooth surfaces aligned with their internal structure in a phenomenon known as cleavage (see e.g., Ford and Dana 1922 Part II.1.277). On the other hand, materials like wood or clay do not necessarily break along smooth faults like those produced by our method. This is a well-studied limitation we share with all mesh-based fracture algorithms and which could be alleviated by borrowing strategies from the literature like the Adaptive Fracture Refinement by Chen et al. [2014], perturbation of crack surface vertices as described by Fan et al. [2022], or the use of pre-authored “splinters” suggested by Parker and O’Brien [2009].

### Table 1. Performance details for all our examples

<table>
<thead>
<tr>
<th>Fig.</th>
<th>#T</th>
<th>Time/mode (s)</th>
<th>k</th>
<th>p</th>
<th>Impact Proj. (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>6316</td>
<td>2.86</td>
<td>30</td>
<td>47</td>
<td>1.06</td>
</tr>
<tr>
<td>18</td>
<td>3931</td>
<td>2.20</td>
<td>20</td>
<td>117</td>
<td>1.21</td>
</tr>
<tr>
<td>14 (a)</td>
<td>3545</td>
<td>0.63</td>
<td>10</td>
<td>35</td>
<td>1.13</td>
</tr>
<tr>
<td>14 (b)</td>
<td>4993</td>
<td>2.63</td>
<td>25</td>
<td>34</td>
<td>1.08</td>
</tr>
<tr>
<td>16</td>
<td>12162</td>
<td>11.6</td>
<td>10</td>
<td>31</td>
<td>1.19</td>
</tr>
<tr>
<td>19</td>
<td>8802</td>
<td>5.91</td>
<td>15</td>
<td>152</td>
<td>1.96</td>
</tr>
</tbody>
</table>

3.7 Choice of strain energy

So far the only requirement on the strain energy density $\Psi$ is that we can construct its second-order approximation near the rest configuration represented by the (positive semi-definite) Hessian matrix $Q$. We now investigate the effect of choices of $Q$ and in particular the relationship with the balancing weight $\omega$.

To make our investigation concrete, take $\Psi$ to be the linear elastic strain energy density, so that $Q$ is the common linear elasticity stiffness matrix. Our observations also follow if one chooses $Q$ to be the Hessian of other, nonlinear energies like the Neo-Hookean or St. Venant-Kirchhoff ones. By sweeping across values of $\omega$ we see a sharp change in the first (and all) fracture mode’s behavior with the discontinuity energy dominating over the strain energy and then sharply switching (see Fig. 12). When the discontinuity energy is effectively zero, then we have simply recovered the usual linear elastic vibration modes (albeit in a convoluted way).

When the strain energy is effectively zero, then we not only start to see sparse fractures, but we also see that each fracture mode undergoes its own zero-strain energy transformation. This behavior is consistent even in larger order modes (inset). That is, each fracture mode undergoes a linearized rigid transformation, the only motions in the null space of the strain energy. Physically, this behavior naturally aligns our fracture modes with the traditional definition of stiff brittle fracture, where materials do not significantly deform before breaking. This is also interesting from a numerical perspective as it implies that the precise choice $Q$ is irrelevant and only its null space matters.

With this in mind, we consider all linearized rigid transformations should be admissible. Since we ultimately care about the fracture pattern created by the modes, we observed qualitatively that the scaling induced by linearized rotations resulted in small elements breaking off and expanding between fragment boundaries to reduce the discontinuity energy. Rather than attempt to identify these as outliers, we found a simpler solution is to work with a strain energy that only admits translational motions in its null space, namely, $\Psi(u, x) = \|\nabla u(x)\|^2$. The Hessian of $\|\nabla u(x)\|^2$ is simply the cotangent Laplacian matrix $L \in \mathbb{R}^{(d+1)(d+1)}$ repeated for each spatial coordinate:

$$Q = I_d \otimes L.$$

This choice of $Q$ is used in all our examples.
Our 1st fracture mode

Fig. 13. Our precomputed fracture modes identify the geometrically weakest regions of a shape, and are activated or not on runtime depending on the nature of the impact.

Efficient precomputation. Our observation regarding the nullspace of $Q$ can be further exploited to greatly reduce the cost of our offline mode precomputation step. Our strain energy being numerically zero in all our modes means all (exploded) vertices belonging to a single element undergo identical deformations. Therefore, by transforming this observation into an assumption, we may store deformations solely at elements, reducing our number of variables by a factor of $d + 1$. This ensures that the strain energy measure on the exploded mesh will always be null, which also means we can remove the quadratic term $u^\top Qu$ from Eq. (16). Further, allowing only per-element deformations also makes our vector discontinuity $D$ necessarily constant along element boundaries, which makes its integral in Eq. (13) trivial without the need of quadrature nodes. The combination of all these observations significantly reduce the size of our conic problem (see Appendix B), allowing computation of identical fracture modes several orders of magnitude faster.

4 TIMING & IMPLEMENTATION DETAILS

We have implemented our main prototype in Python, using Libigl [Jacobson et al. 2018]. We used Mosek [ApS 2019] to solve the conic problem in Eq. (16). We report timings conducted on a 2020 13-inch MacBook Pro with 16 GB memory and 2.3 GHz Quad-Core Intel Core i7 processor. To produce our animations, we follow a traditional Houdini [SideFX 2020] fracture simulation workflow, exchanging the usual Voronoi or openVDB fracture nodes for our own fractured meshes. Our impact projection step could be fully integrated into Houdini’s rigid body simulator at a minimal performance cost. Only for simplicity in prototyping, we chose not to do this and instead compute our final mesh $Ω^*$ in Python taking into account the animation’s impact and load it into Houdini directly to show a prototype of what our algorithm would look like integrated in a rigid body simulation.

Our algorithm’s only parameters are the tolerances $\varepsilon$ and $\sigma$, which we fix at $\varepsilon = 10^{-10}$ and $\sigma = 10^{-3}$. As we discuss in Section 3.7, a scalar $\omega$ will not actually have an effect in the output as long as it is small enough for us to be within the zero-deformation fracture realm (see Fig. 12).

Our proposed algorithm works in two steps. First, we precompute a given shape’s fracture modes. This step takes place offline, following Algorithm 1. The computational bottleneck of this section of our algorithm is the conic solve detailed in Eq. (16). Each mode takes between 0.5 and 12 seconds to compute in our meshes, which have between 3,000 and 15,000 tetrahedra.

Secondly, our impact projection step as detailed in Section 3.3 is the only part of our algorithm that happens at runtime. The complexity of this step is dominated by the projection step in Eq. (20), which is $O(kn)$, where $k$ is the number of precomputed modes and $n$ is the number of vertices in the boundary of the connected components described in Section 3.4. All other elements of our projection step are $O(p)$, where $p$ is the number of connected components (in our example between 10 and 500) and $p << n$ so they can be disregarded from the complexity discussion. In our examples, $n$ is between 1,000 and 10,000 and we compute between $k = 20$ and $k = 40$ fracture modes, meaning our full runtime step requires between 0.1 and 1 million floating point operations, putting it well within realtime requirements, even if one greatly increases $n$, $k$ or the number of objects on scene (note this projection step only needs to be run when a collision is detected, and not at every simulation frame).

Our unoptimized, CPU implementation takes between one and two milliseconds to carry out this step on our laptop (see Table 1).
5 EXPERIMENTS & COMPARISONS

Our proposed fracture modes naturally identify the regions of a shape that are geometrically weak, as opposed to existing procedural prefracture algorithms. We make this explicitly clear in Figs. 3 and 15, where existing prefracture work fails to identify even the most obvious intuitive breaking patterns which are present in our first (non-trivial) fracture mode. Even in less didactic examples, Voronoi-based prefracture methods result in convex, unrealistic and easily recognizable pieces (see Figs. 22 and 2), while our fracture modes are realistic and can produce a much wider set of shapes.

"Realism" in a fracture simulation is a hard quantity to evaluate; however, there exist works on structural analysis like [Zhou et al. 2013] that identity the weakest regions of a given object. In Fig. 5, we show how our fracture modes produce breaking patterns that align both with their analysis as well as with their real-life experiments.

We model heterogeneous and anisotropic materials by incorporating a vector field \( \eta : \Omega \rightarrow \mathbb{R}^d \) to the discontinuity energy:

\[
E_D(u) = \sum_{i=1}^{p} \sqrt{\int_{S_i} \| \eta(x) \circ D(u, x) \|^2 dx},
\]

where \( \circ \) denotes Hadamard (elementwise) multiplication. In Fig. 18, we experiment with varying the magnitude of \( \eta \) as an artist control tool to designate regions that should not fracture. In Fig. 17, we make \( \eta = (10, 10, 1) \) to favour vertical faults over horizontal ones.

5.1 Fracture simulations

Our proposed method is ideal for use in interactive applications. In Fig. 4, we show screenshots of our 2D realtime fracture interactive app. The user can cause different impacts on the object and see the fracture patterns that result from projecting them onto our modes.

The interactive Computer Graphics application par excellence is video games. In Fig. 7, we show a prototype where our precomputed fracture modes for a Space Wizard Vehicle can be stored so that the player sees different fracture behaviours depending on the received impact. In Fig. 14, we precompute the fracture modes for two different vehicles and show how they break under a similar impact.

Our algorithm can be used for any realtime fracture application, from simple objects breaking into solid pieces in the foreground of an animation (see Fig. 13) to thin shells shattering upon impact (see Figs. 19 and 18). In Fig. 16, we use our fracture modes to simulate a human skull breaking into many pieces upon impact with a wall.

6 LIMITATIONS & FUTURE WORK

Our fracture modes method is intended for stiff brittle fracture. We conjecture that general rigid fracture and even ductile fracture simulation could also benefit from our sparse-norm formulation. In future work, we would like to improve the performance of our precomputation optimization. We experimented with Manopt [Boumal et al. 2014], but so far observed significantly slower performance than our proposed method. For very large meshes, the projection step could exceed CPU usage allowances for realtime applications. It may be possible to conduct this entirely on the GPU.

Our fracture modes are global in nature, meaning they create relations between regions of the object that will not typically fracture together (unlike other prefracture methods like [Oh et al. 2012]). A way of preventing a fracture in one location from also causing an undesired fracture elsewhere is to use our modes only to identify the pieces that could break off of an object in the precomputation step, and swap our realtime impact projection for a least squares constant-per-piece approximation (see Fig. 21).

Our use of an exploded mesh \( \tilde{\Omega} \) allows us to expand the usual finite element hat function basis to include discontinuities along element boundaries. This mesh dependency is not present in traditional Voronoi or plane-cutting prefracture algorithms, and can
lead to visible artifacts if the simulation mesh is too coarse. We alleviate it with post-facto smoothing (Section 3.6). Another way of reducing it (at a performance cost) would have been to include basis functions with sub-mesh-resolution discontinuities in the style of XFEM [Chitalu et al. 2020; Kaufmann et al. 2009].

Our algorithm is designed to fit into realtime rigid body simulations like those encountered in video games. Thus, our outputs will not contain partial fractures (unlike e.g., [Müller et al. 2013]).

Secondary fractures were not included in our simulations. Computing a new set of fracture modes for each piece would exceed realtime constraints. While one could obtain plausible secondary fractures by restricting our precomputed fracture modes and pattern to each primary fracture piece (see inset), there is no guarantee that these would match the individual piece’s fracture modes.

We use a vector field $\eta$ to model an object’s preferred breaking behaviour (see Figs. 18 and 17). Promising future work includes incorporating physical material properties into our mode computation; for example, by treating tangential and normal discontinuities differently.

Our method belongs to the class of prefracture, not dynamic algorithms. Nonetheless, our method can be evaluated on dynamic fracture benchmarks like the notched block in [O’Brien and Hodgins 1999]: if a given fracture plane is contained in one of our fracture modes, it can be present in the fractured output (see Fig. 20). The fracture fault will be the same regardless of the directionality of the impact. This deviates from the real-world mechanical behaviour, where faults will be different for brittle materials under uniaxial tension, pure shear, and torsion loads (see [Lawn 1993], Chap. 2).

We hope our introduction of fracture eigenmodes inspires the realtime simulation community further to use the well-studied tools of modal analysis to this rich problem, and the broader Computer Graphics research community to look at other open problems with this modal lens.

ACKNOWLEDGMENTS

This project is funded in part by NSERC Discovery (RGPIN2017–05235, RGPAS–2017–507938), New Frontiers of Research Fund (NFRF–201), the Ontario Early Research Award program, the Canada Research Chairs Program, the Fields Centre for Quantitative Analysis and Modelling and gifts by Adobe. The first author is supported by an NSERC Vanier Scholarship and an Adobe Fellowship. The four middle authors were supported by the 2020 Fields Undergraduate Summer Research Program.

We acknowledge the authors of the 3D models used throughout this paper: MakerBot (Fig. 1, CC BY 4.0), HQ3DMOD (Figs. 6 and 19, TurboSquid 3D Standard Model License), Freme Minskib (Fig. 7, CC BY-NC 4.0), 3Demon (Fig. 9, CC BY-NC-SA 4.0), Reality_3D (Fig. 11, CC BY 4.0), Alex (Fig. 14, CC BY-NC-SA 4.0), Falha Tecnologia (Fig. 18, TurboSquid 3D Standard Model License), LeFabShop (Fig. 16, CC BY-NC 4.0), The Database Center for Life Science (Fig. 15, CC BY 4.0), and Gips (inset in Section 3.5, CC BY-NC 4.0).

We are grateful to the anonymous peer reviewers for their insightful suggestions. We would especially like to thank Reviewer #3 for inspiring the shockwave-based impact projection in Section 3.3. We would also like to thank Chris Wojtan, David Hahn and Klint Qinami for early experiments and discussions of sparse-norm fracture models; Eitan Grinspun, David I.W. Levin, Oded Stein and Jackson Phillips for insightful conversations; Rinat Abdashitov for providing an implementation of his algorithm mentioned in Section 3.6; Qingnan Zhou for providing the 3D models used in Fig. 5; Xuan Dam, John Hancock and all the University of Toronto DCS staff that kept our lab running during the hardest of times.
REFERENCES


A CANONICAL CONIC PROGRAM FORM OF EQ. (16) Let us define a sparse matrix $D \in \mathbb{R}^{p \times (d+1)m}$ that operates on a deformation map and evaluates the vector-valued discontinuities at all the relevant integration quadrature points. The ordering of the rows of $D$ is arbitrary, and we choose it such that $D$ can be separated into $d$ blocks, one for each quadrature point.

For example, in the case $d = 2$, we choose said ordering such that the first $p$ rows of $D$ are the vector-valued edge-wise discontinuity

$$\sqrt{\frac{2}{d}} \left(\frac{1}{\sqrt{3}}\right)$$

from Eq. (13) and the last $p$ rows of $D$ correspond to
Since $Q$ is positive semi-definite, we can write it as $Q = R^T R$ for some matrix $R$. Define

$$Y_e = (Du)_e, \forall e = 1, \ldots, 2p.$$  \hfill (27)

where $l_e$ is the length of edge $e$. Next, define

$$r_i = (Ru)_i, \forall i = 1, \ldots, d(d + 1)m.$$  \hfill (28)

Then, Eq. (16) can be written in the canonical form

$$\begin{array}{c}
\arg\min_{u, t, Y, z} \\
\begin{bmatrix}
1 & t^T \\
0 & r \\
0 & u \\
0 & Y \\
1 & z
\end{bmatrix}
\end{array}
\begin{bmatrix}
0 \\
0 \\
0 \\
1 \\
1
\end{bmatrix}
$$  \hfill (29)

subject to

$$t \geq \sqrt{r_1^2 + \ldots + r_{d(d+1)m}^2}$$

$$z_e \geq \sqrt{\sum_{s=0}^{d-1} \|Y_{e+sp}\|_2^2} \forall e = 1, \ldots, p$$

$$Y = Du$$

$$r = Ru$$

$$c^T \tilde{M} u = 1$$

$$U^T C^T M v = 0, \quad \forall j = 1, \ldots, i - 1.$$  \hfill (32)

B EFFICIENT CONIC PROGRAM FROM SECTION 3.7

Let $C$ be the $\mathbb{R}^{(d+1)mxm}$ matrix of ones and zeros that transfers values from elements to vertices in the exploded mesh, and let us assume now that we are storing per-element deformations in vectors $v, U_j, c \in \mathbb{R}^{dm}$. Then, our conic problem from Appendix A becomes

$$\begin{array}{c}
\arg\min_{v, t, Y, z} \\
\begin{bmatrix}
1 & t^T \\
0 & r \\
0 & v \\
0 & Y \\
1 & z
\end{bmatrix}
\end{array}
\begin{bmatrix}
0 \\
0 \\
0 \\
1 \\
1
\end{bmatrix}
$$  \hfill (30)

subject to

$$t \geq \sqrt{r_1^2 + \ldots + r_{d(d+1)m}^2}$$

$$z_e \geq \sqrt{\sum_{s=0}^{d-1} \|Y_{e+sp}\|_2^2} \forall e = 1, \ldots, p$$

$$Y = DCv$$

$$r = RCv$$

$$c^T \tilde{M} u = 1$$

$$U^T C^T M v = 0, \quad \forall j = 1, \ldots, i - 1.$$  \hfill (32)

By construction, $QCv = 0 \Rightarrow RCv = 0$, which means we can remove $t$ and $r$ as variables entirely. Further, since the vector-valued discontinuity is constant across element boundaries, $Y_e = Y_{e+sp}$, with $s = 0, \ldots, d - 1$, meaning that we can remove the summation from the $z_e$ cone and consider only the first $p$ rows of $Y$. This leads to the equivalent, simpler conic program

$$\begin{array}{c}
\arg\min_{v, t, Y, z} \\
\begin{bmatrix}
0^T \\
0 \\
0 \\
1
\end{bmatrix}
\end{array}
\begin{bmatrix}
v \\
Y \\
z
\end{bmatrix}
$$  \hfill (31)

subject to

$$z_e \geq \sqrt{d \|Y_e\|_2^2} \forall e = 1, \ldots, p$$

$$(Y)_e = (DCv)_e \forall e = 1, \ldots, p$$

$$c^T C^T M v = 1$$

$$U^T C^T M v = 0, \quad \forall j = 1, \ldots, i - 1.$$  \hfill (32)