Workshop on Virtual Reality Interaction and Physical Simulation VRIPHYS (2018)
F. Jaillet, G. Zachmann, K. Erleben, and S. Andrews (Editors)

Quantitative Validation of Physically Based Deformable Models in Computer Graphics

M.K. Banks[1], A.L. Hazel[2] and G.D. Riley[1]

[1]School of Computer Science, The University of Manchester, U.K.
[2]School of Mathematics, The University of Manchester, U.K.

Abstract

This paper presents a novel software framework for quantitative validation of physically-based deformable models in computer graphics. In the majority of previous studies, validation is qualitative (through visual plausibility), which is necessarily user subjective. The proposed framework facilitates construction of a single scalar output that quantifies the agreement between the complete time histories of test and reference models. Different models and comparison metrics can be easily included within the general framework. The framework is shown to yield a high accuracy score for a simplified model that can be analytically derived from the reference model, indicating that the framework is reliable. A lower score results when evaluating a more approximate, yet still visually plausible, model, demonstrating the objective sensitivity of the framework. The software framework can thus provide an objective measure of accuracy and a standardised way to quantitatively compare the accuracy of one method against another, whilst also supplying a quantitative rationale for trading accuracy and performance.

CCS Concepts

•***Computing methodologies*** → ***Model verification and validation;*** *Physical simulation; Simulation evaluation;*

1 Introduction

The increase in power of modern day processors means that it is now possible to simulate physically based deformable models (PBDMs) within computer graphics (CG) that perform well (often in real-time) and are accurate enough so as to be realistic, all whilst maintaining their stability. A high level of accuracy is important as it helps with the immersion and interactions of the user into the virtual environment. Accuracy, performance and stability have contrasting computational requirements. A more accurate model requires greater computational resources, whereas if performance is important then some degree of accuracy will typically have to be sacrificed to accommodate the performance improvements.

Our contribution in this work is a software framework that can be used to objectively quantify the accuracy of competing PBDMs used in CG. It quantifies the agreement of the time-evolving deformations of a test PBDM against a reference PBDM's deformation in a selected scenario through a single scalar output quantity. Whilst this quantity is not a diagnostic tool to identify sources of errors, to the best of our current knowledge obtaining a single quantity for accuracy over a prescribed simulation time is not something that has previously been studied. The framework is a component-based software architecture whose necessary components are replaceable and connected to produce the output quantity in the range $[0,1]$, with a score of 1 being a maximal score of agreement between series. An application of this framework is the ability to directly compare the outputs produced by different test PBDMs to establish a ranking of accuracy.

In virtual reality applications, a user's interaction with entities in the virtual environment improves immersion [NLB*07]. In this interaction, stability and performance of a PBDM's simulation hold more importance than its accuracy [NMK*06]. Some non-physically based models have been developed that knowingly sacrifice accuracy for the sake of performance gains [MHTG05]. On the other hand, in surgical simulation, accuracy is understandably of the utmost importance [ZZGC17]. Indeed, the order of priority of performance, accuracy and stability is very much dependent on the specific field within CG being targeted. Each PBDM proposed in the literature balances the three factors differently and so is more suited to some fields than others.

The quantification of accuracy is explored through the procedure of verification and validation (V&V). Whilst ultimately dependent on the intended use of the application, a definition of V&V is given in [OT02] as:

1. "*Verification*: the process of determining that a model implementation accurately represents the developer's conceptual description of the model and the solution to the model."
2. "*Validation*: the process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model."

DOI: 10.2312/vriphys.20181062

Verification might be viewed as the successful implementation of the model on a computer, whilst validation is whether or not that computational model's solutions hold true to its real-life or analytical counterparts. Indeed, "the ultimate test of accuracy is how well the numerical method reproduces known results" [Mon92].

The structure of the paper is as follows. Section 2 looks at related work in the literature. The design and a description of the framework is given in Section 3. As example applications of the framework, two popular PBDMs from the literature are evaluated in a familiar cantilevered beam scenario in Section 4. Both PBDMs are mass-spring systems (MSS): a collection of discrete point masses connected by Hookean springs. The two PBDMs differ in their approach to solving the equations of motion, with the latter being a more approximate solution than the former that provides greater performance. The results, also presented in Section 4, are analysed and their objective scores of accuracy directly compared. Conclusions and a discussion of the limitations and future work complete the paper in Section 5.

2 Related Work

It is common for validation of PBDMs in CG to be only qualitative, as has been expressed in surveys [NMK*06] [MBP14]. Often the acceptance criteria for accuracy is through *visual plausibility* [BMO*14] [NMK*06]. A disadvantage of qualitative validation is that it is subjective; the validity of a PBDM is dependent on the user's senses, which are not always reliable. It is difficult to detect subtle abnormalities in a simulation [BET14] and reliable sensory validation requires full concentration throughout [NLB*07]. However, the qualitative validation of a subject matter expert (SME) [BFZ*13] in the field must also retain its significance and cannot be disregarded because it is qualitative. Thus, any quantitative validation ought to be consistent with the qualitative validation of any such SME [SKH*10]. Nevertheless, there is a recognition of room for more concrete assessment [GM97] [MBP14] and the need to validate results with the real world via a standard set of scenarios is acknowledged [MBP14] [SSP07]. The finite element method (FEM) at a suitably high resolution is regarded as being the most accurate PBDM when used to provide a solution to nonlinear solid mechanics problems (NLFEM) [MM07] [FWY] [Glo12]. The determination of a sufficiently high resolution can be automated by using *a posteriori* error estimates [HH06] [ZZ92]. If the estimated error is deemed too large in a particular region then the mesh can be refined more finely in this area.

In interactive CG, quantitative validation techniques have seldom been used. In evaluating the material properties in a PBDM, both the Hausdorff distance [HKR93] and mean absolute error (MAE) of displacements of points on the body are used to validate the results of stretching a cloth. Real-life test data has been compared with a NLFEM model [BKCW14]. These measures are taken at the end of the stretching (only) and so, being a snapshot, serve as a static evaluation of accuracy. The history of the deformation is not measured. To compare the time-evolution of such measurements would evaluate the dynamic, temporal agreement; one can imagine a scenario whereby the end deformations are in good agreement even though the deformation histories are not. A field in engineering that assesses the temporal agreement of reference and test data is vehicle safety, which is discussed in section 2.3.

2.1 Quantitative Validation

The establishment of quantitative validation requires the evaluation of some quantities to show agreement. Sarin *et al.* define an *error measure* as a quantity associated with differences of a particular feature of a time series, whereas an *error metric* is an overall value of discrepancy between time series [SKH*10]. An error measure will refer to an isolated quantification of agreement that alone is not suitable for quantitative validation. This includes a quantification of agreement at a single time instance i.e. the static evaluation mentioned previously. An error metric will reflect the agreement of the entire (deformation) histories. If *uncertainties* in data acquisition are considered, a *stochastic* validation metric is obtained, otherwise it is *deterministic* [SKH*10] [LCAH11]. Only deterministic validation metrics are considered in this work.

Measured errors in numerical computation can arise due to both model and discretisation error [BTC*17]. Model error is the error induced by the mathematical formulation of a physical problem and is quantified in the framework presented here. Discretisation error is the error arising due to the discretisation (approximation) of a mesh that is modelling a continuous volume. An increase in resolution might lead to the convergence of a numerical solution; this is the minimisation of discretisation error [BTC*17]. Error estimates work to minimise the discretisation error in finite element solutions [HH06]. Once minimised, any other error against some ground truth is primarily due to modelling error. Validation metrics that quantify discretisation error are deterministic because the experimental data sets are being compared with certainty (even though the data sets may have been obtained stochastically) [SM15], so the discretisation error is minimised deterministically.

2.2 Validation in Surgical Simulations

Validation in surgical simulation comprises deterministic validation measures and the evaluation of data at a single point in time [PIS*13] [LYD*12] [YYT*15] [CAK*14] [YYT*15] [CAK*14] [HDBC14] [DHC*13] [DHC*16] [MADC08] [W*12]. Computational simulation data is compared against real-life test data and then verified based on a chosen end state. The error measure operates on displacements of points between data sets and is typically a central tendancy (e.g. the mean value) and variation [PIS*13] [LYD*12] [YYT*15] [CAK*14], maximum error [YYT*15] [CAK*14] or some variation of the Hausdorff distance [HDBC14] [DHC*13] [DHC*16]. This overwhelming focus on end-state validation ignores the histories of deformations. Nevertheless, test data exists in the form of the so-called *Truth Cube* for real-time soft tissue deformation [KCO*03] [ZZGC17]. The test data is readily available as nodal positions for a series of quasi-static loads.

In evaluating the trade-off between accuracy and performance of a PBDM, the plotting of root-mean-square error (RMSE) (of nodal displacements of the test model against a reference model) against time displays a trade-off between accuracy and performance [CAK*14]. Whilst this allows for visual, qualitative confirmation, there is no translation of this error measure history into a

quantity that might be used to compare two time series objectively, as we do here.

A step needs to be made to transform a collection of error measures over time into an error metric. As mentioned at the start of section 2, any sort of discrepancy between an entire pair of time series that gives rise to some kind of quantification metric should also reflect the views of a SME. As an example use of SMEs' validation, a survey is conducted by Sui *et al.*. A deformation - modelled using a popular PBDM called position-based dynamics (PBD) [BMO*14] - is evaluated by surgeons and produces results of around 90% mean satisfaction with a $5-10\%$ standard deviation [SPQ*17]. A high mean score with a comparatively small spread, sourced from SMEs, provides very good validation that lies somewhere between qualitative and quantitative. The ratings from SMEs are the most informed and reliable, and should be the most objective. As such, error metrics that are developed to reflect SMEs' opinions should be used. This is precisely what is done in the engineering field of vehicle safety.

2.3 Validation in Vehicle Safety

Qualitative validation should be based on the knowledge of subject matter experts (SMEs) [BFZ*13]. A problem with this is that each design iteration - be it large or small - requires the constant cooperation of developers and SMEs. This can be a slow back and forth conversation that can hinder productivity should the SME not be readily available.

In vehicle safety, some signals, plotted as curves on a graph, are quantitatively compared and scored to show the level of agreement between them. The purpose is to compare two unambiguous signals; some measure is plotted against time, once for (real-life) test data serving as the reference and once for computational data serving as the comparison [SKH*10] [Wei17], with the evaluation reflecting that of SMEs. Since the field is vehicle safety, the tuning of parameters is strict so as to comply with the obvious safety constraints that such a field demands.

Two metrics that quantify the agreement between pairs of signals are the CORrelation and Analysis (CORA) metric [GGW09] and the Enhanced Error Assessment of Response Time Histories (EEARTH) metric [BFZ*13]. CORA uses two independent sub-ratings. The first is a corridor rating that assesses, at each sample point of the signal, how close the corresponding point on the comparison curve is through the use of "corridors" around the reference curve. The second is a cross correlation rating that directly assess the curves' characteristics such as magnitudes, shapes and phase shifts. The idea is that each sub-rating makes up for the disadvantages of the other. EEARTH, as the name suggest, enhances EARTH that uses three sub-ratings relating to phase, magnitude and curve topology [SKH*10]. Parameters for both CORA and EARTH are tuned by the user for the application. EEARTH has its parameters either fixed by SMEs or automatically calculated depending on the reference curve and produces a score in the range $[0,1]$, with 1 being the best score for matching signals.

The ISO18571 standard seeks to develop an error metric that utilises the best features from CORA and EEARTH, the leading metrics in the field [BFZ*13]. EEARTH's magnitude (e_M), phase (e_P) and topological (e_T) error ratings and CORA's corridor rating (e_C) form a weighted linear combination to give

$$e_{\text{ISO18571}} = 0.4e_C + 0.2e_M + 0.2e_P + 0.2e_T \in [0,1]. \quad (1)$$

The following thresholds are given by SMEs and remain valid only when the parameters of the individual error measures are unchanged from the ISO standard [BFZ*13] [VDSG15] [SWU*15]:

Rating, R	Grade
$0.94 < R \leq 1.00$	Excellent
$0.80 < R \leq 0.94$	Good
$0.58 < R \leq 0.80$	Fair
$0.00 \leq R \leq 0.58$	Poor

The ISO18571 standard is used in the framework in Section 3.2 and the above error metric grades will be used in the results of Section 4 to describe the accuracy of the tested PBDMs.

3 Framework

The strategy for quantitative validation requires selection of error measures and error metrics, as defined in section 2.1. It also requires the selection of scenarios to simulate in which the errors can be evaluated. The design of the framework is presented in this Section together with a description of its parts.

The selection and specification of error measures, error metrics, scenarios and PBDMs is made using components. These components are modular, architectural units; they are primary computational elements and data stores of the system with well-defined interfaces that expose their functionalities [LW07]. The software architecture for the framework's components is presented in Figure 1 as a component diagram, modelled in UML 2.0 [PP05]. The ports, representing the interfaces, serve as either sinks or sources for receiving or sending data, respectively (these are "sockets" and "lollipops" in UML 2.0), and are connected via assembly or delegation [LW07]. The components are composed in the design phase of the component life cycle; all component compositions are made in the design stage and persist until the runtime with no further amendments [LW07]. The user can substitute appropriate components without changing the framework's functionality.

An Entity component is composed of a Body component comprising a set of nodes (or "physics state", for simulation) and a Mesh component for regular rendering of those nodes as vertices (or "render state"). Thus the mesh component depends on the Body component's physics state. The deformation of the body's nodes is governed by a Deformation Model component (Section 3.4) through an assembly connection on the *deform* interface. The body requires initialisation, which is provided through delegation connections to the Scenario component (Section 3.3) by the Entity component and then the Supervisee component. The Supervisee component contains two Entity components; one serves as the reference entity, whose state is queried against by the second, test entity. As the name suggests, the Supervisee component is supervised in the framework by the encompassing subsystem (not shown), driving the simulation and rendering the states. It also contains two DTS (discrete time series) components that each require data and provide a series. The required data is provided, through delegation by the Supervisee component, by an Error Measure component (Section 3.1). This Error Measure component itself requires

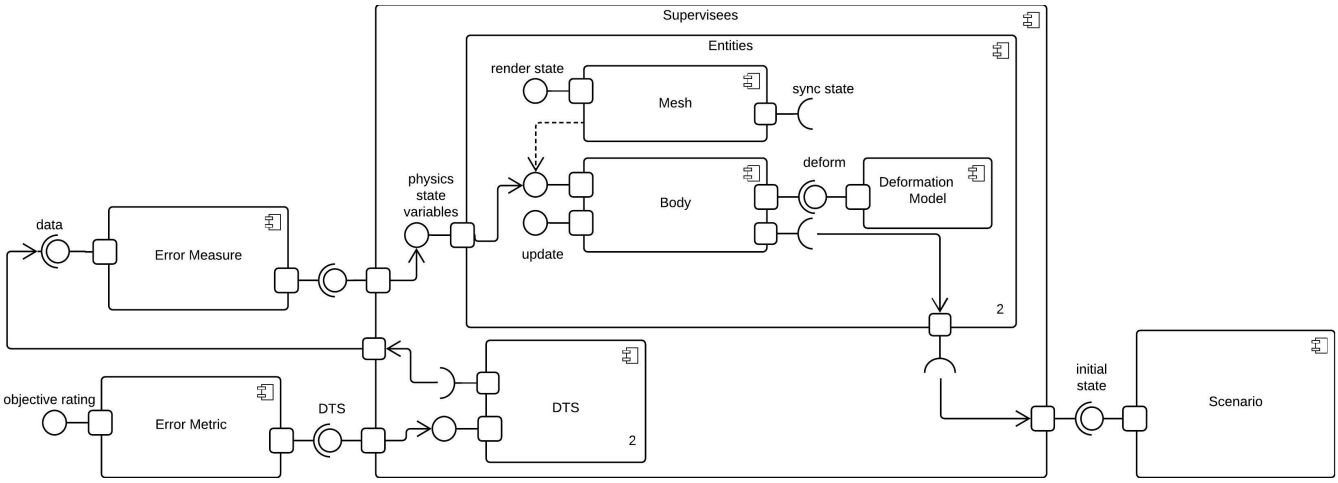

Figure 1: *A UML 2.0 component diagram of the software framework.*

a physics state, realised on the Body component (through delegation connectors on the Supervisee and Entity components). Finally, the provided time series from the DTS components are passed to an Error Metric component which produces an output rating (Section 3.2). Components may be substituted provided, of course, that the contracts set forth by the interfaces are abided by. Each component in the current framework is explained further below. The PBDM that is being evaluated is inserted in as the Deformation Model component on the test Entity component.

3.1 Error Measure Component

The error measure measures some quantity of the state of a PBDM at a single instance in time, relative to some reference state. The reference state will be the initial positions of the bodies in that particular scenario. As discussed in section 2.2, variations of the Hausdorff distance are popular choices for error measures in the literature when some shape matching algorithm is used. The directed Hausdorff distance is defined as [HKR93]:

$$d_{DH} = \max_{a \in A} \min_{b \in B} ||a - b||, \tag{2}$$

where A and B are sets of positions in Euclidean space and so $||\cdot||$ is the Euclidean norm between points $a \in A$ and $b \in B$. This measure is the greatest of all distances from each point $a \in A$ to the closest point $b \in B$ and can be thought of here as how much the test and reference bodies overlap (a measure of 0 is a perfect overlap). Due to its popularity in shape matching, the bidirectional Hausdorff distance [HKR93] is used in the framework:

$$d_{BDH} = \max(d_{DH}(A,B), d_{DH}(B,A)). \tag{3}$$

Alternative choices are possible and require a simple substitution of the Error Measure component. For example the modified Hausdorff distance [DJ94] in place of (2) in (3), which can overcome the influence of outliers by taking an average:

$$d_{MDH} = \frac{1}{N_a} \sum a \in A \min_{b \in B} ||a - b||. \tag{4}$$

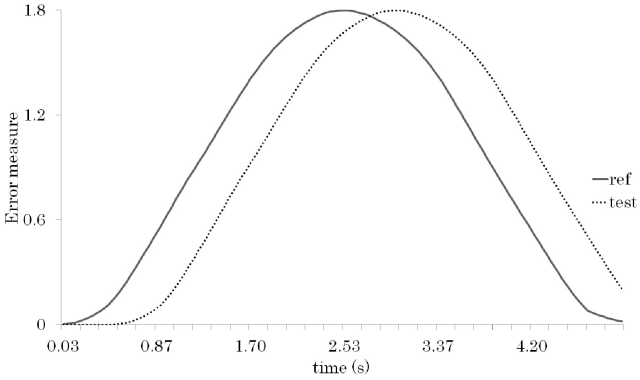

Figure 2: *An example test and reference signal pair over five seconds.*

3.2 Error Metric Component

The error measures outlined in Section 3.1 are plotted as discrete time series serving as signals to be used in $e_{\text{ISO}18571}$ (as described in Section 2.3). A score for the test scenario is produced and serves as the measure of accuracy. The error metric $e_{ISO}18571$ measures how close the test curve is to the reference curve, whilst also taking into account the magnitude, phase and slope similarities. A set of signals might look like Figure 2. The test signal is delayed to start $0.5s$ later but is otherwise the same as the reference signal. The metric score of $e_{ISO}18571 = 0.711$ reflects the signals' similarities and accounts for the phase discrepancy independently of the other signal properties under evaluation by the error metric.

3.3 Scenario Component

The selection of scenarios has to address the following factors:

1. Test data: does either an (exact) analytical solution or experi-

mental data exist against which we can compare our simulated results?
2. Complexity: are typical deformations covered? Are the scenarios sufficiently complex enough (i.e. nonlinear deformations) to reflect real-life deformations?
3. Geometric simplicity: are the continua easy to model in a PBDM? Can a mesh be easily constructed?
4. Duration: over what timescale should the simulation remain accurate?

3.4 Deformable Model Components

There are two PBDMs that must be provided. The first is a test PBDM that is under evaluation within the framework. As stated, this test Deformation Component can govern the deformation of the test Body Component's nodes (otherwise the Body does not deform). The user provides an implemented test Deformation Component in the framework that is to be evaluated. The second PBDM serves as the reference and can be an analytic solution, should one exist in the particular scenario. In scenarios for which no analytic or test data exists, a suitably accurate PBDM can be used as reference instead. This suitability can be determined by an ISO18571 error metric result that grades as "excellent" (see Section 2.3) in scenarios for which there is test data available. The NLFEM provides an approximate solution to the equations of motion from continuum mechanics (the principle of virtual displacements [HH06]), with a St. Venant-Kirchhoff strain energy function. It is used in the literature as a reference model for validation and is used in the framework here. Its suitability as a reference PBDM is demonstrated in Section 4.1.

4 Results

The framework is used to undertake quantitative evaluation of two popular example PBDMs from the literature (Section 4.2). In Section 4.1.1, the NLFEM is shown to be a suitable reference PBDM, whilst a particular test scenario is outlined in Section 4.1.2 that will be used in the evaluations of Section 4.2. Analysis of the results of the evaluations is given in Section 4.3.

4.1 Test Scenarios

As mentioned in Section 3.4, if neither an analytic solution nor a set of test data exist for a test scenario, a suitably accurate PBDM should be used as the reference PBDM. Thus, the NLFEM is evaluated in the framework in a steady state 2D cantilever beam scenario - for which an approximate analytic solution can be derived - for different levels of discretisation, in order to demonstrate the high accuracy of the NLFEM and to justify its use as a suitable reference model. Following this, an unsteady 3D cantilever beam scenario is presented that adheres to the scenario requirements of Section 3.3, together with the NLFEM configuration that will serve as the reference PBDM. This scenario and reference PBDM combination will be used in evaluating test PBDMs in Section 4.2.

4.1.1 Establishing a gold standard: Steady 2D Cantilever Beam

The classical solid mechanics problem of the bending of a cantilever beam subject to a constant pressure loading, P, on its upper face is as in Figure 3. The balance of linear momentum is given by

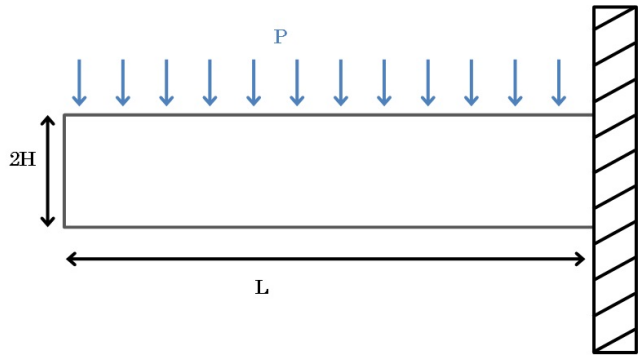

Figure 3: *The setup of the cantilever scenario.*

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \frac{\partial \tau_{ij}}{\partial x_j} + F_i, \tag{5}$$

where ρ is the continuum's density, F is the body force, u is the displacement, x is a coordinate direction and τ is the second Piola-Kirchhoff stress tensor. An approximate (steady state) solution for the stress field can be constructed from the following Airy stress function, valid in the absence of body forces ($\frac{\partial^2 u_i}{\partial t^2} = F_i = 0$) and for small pressures [HH06]:

$$\Phi(x,y) = ax^2 + bx^2 y + c(x^2 y^3 - \frac{1}{5}y^5) + dy^3, \tag{6}$$

where

$$a = -\frac{1}{4}P, b = -\frac{3}{8}\frac{P}{H}, c = \frac{1}{8}\frac{P}{H^3}, d = \frac{1}{20}\frac{P}{H}.$$

The displacement field can be recovered as (μ is a Lamé constant, derivable from E and ν):

$$\begin{aligned} u(x,y) = \frac{1}{2\mu}\Big(&(1-\nu)\big(2ax + 2by(x-L) \\ &+ 2cy(x^3 - xy^2 - L^3) + 6dy(x-L)\big) \\ &+ 2a(\nu L - x) - 2by(x+L) - 2cxy^3\Big); \end{aligned} \tag{7}$$

$$\begin{aligned} v(x,y) = \frac{1}{2\mu}\Big(&(1-\nu)\big(2ay + b(y^2 - x^2 + 2Lx - L^2) \\ &+ c(3x^2 y^2 - \frac{1}{2}(y^4 + x^4 + 3L^4) + 2L^3 x) \\ &+ 3d(y^2 - x^2 + 2Lx - L^2)\big) \\ &+ b(3L^2 - x^2 - 2Lx) + c(y^4 - 3x^2 y^2) \\ &+ 3d(2Lx - x^2 - y^2 - L^2)\Big). \end{aligned}$$

The pressure is incremented at each timestep to give a series of quasi-static loads. The length is set to $L = 5$ and the height to $2H = 1$ so as to non-dimensionalise all lengths. A NLFEM solution is computed using *oomph-lib* [HH06], with $E = 250\text{Nm}^{-2}, \nu = \frac{1}{3}$. This is compared against the analytic solution, serving as the reference. To eliminate discretisation errors in the NLFEM model, the error metric rating of Section 3.2 is calculated for different discretisations in the beam's body. Figure 4 shows this. The error metric achieves a score of at least 0.97 - excellent by ISO18571 - for discretisations of $x \geq 15, y \geq 3$ nodes and finer. Whilst mesh refinement proce-

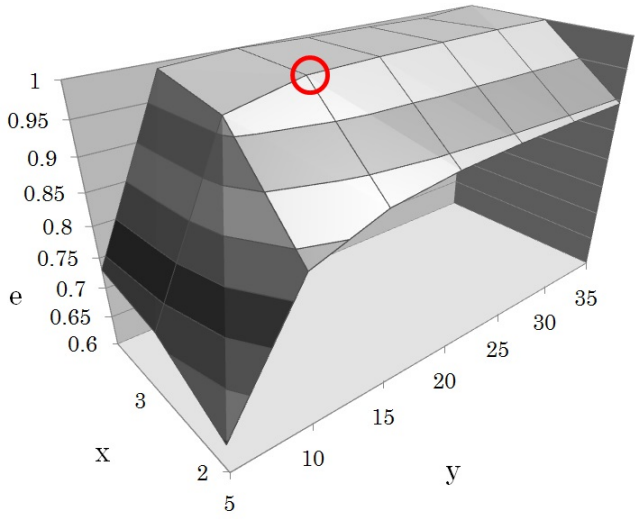

Figure 4: *A graph showing the ISO18571 error score, e, for different discretisations of the 2D cantilever in x and y. The point circled in red is the minimum discretisation for a score of at least 0.97.*

dures using *a posteriori* error estimates can (and will) be used in scenarios for reference model calculations, here the sensitivity of the framework to a purposeful increase in mesh resolution can be observed through the corresponding error metric scores in Figure 4. The high scoring error metric scores validate NLFEM to be used as a reference in future evaluations in the framework.

4.1.2 Unsteady 3D Cantilever Beam

The geometry is the same as 4.1.1 except it is extended to three dimensions. The inertia term is included in (5), whilst the pressure loading is removed. Instead, a body force of $(0,-1,0)^T$ is applied. This has the effect of the beam being previously steady in its initial position and instantaneously being subjected to the body force at the start of the simulation. The body is perfectly elastic; no damping occurs and the body oscillates indefinitely. The approximate analytic solution, (7), is for steady state only and is therefore no longer valid. The undeformed body is a parallelepiped shape and simple to recreate. The simulation is run for 5 seconds to allow for a significant deformation that generates an oscillation. The cantilever is $4\times1\times1\text{m}^3$ and is both homogeneous and isotropic, with the elastic modulus, $E = 400\text{Nm}^{-2}$ and Poisson's ratio, $\nu = \frac{1}{4}$. At fixed time steps, dt, the simulation is advanced by dt. The Hausdorff distance of the beam to its initial position is calculated at each time step. The reference NLFEM, serving as the gold standard, is discretised with the use of the Z^2 error estimator [HH06] [ZZ92] with strict error thresholds (minimum and maximum permitted errors of order 10^{-4} and 10^{-3}, respectively [HH06]) to give a non-uniformly discretised mesh of greater than 4000 vertices. Different, unrefined and uniform resolutions provide solutions that are compared with this gold standard in Figure 5. A particularly noteworthy observation is the error metric's sensitivity to the convergence towards the gold standard. It remains to show what the minimum simulation updates per second can be without sacrificing accuracy. Whilst this

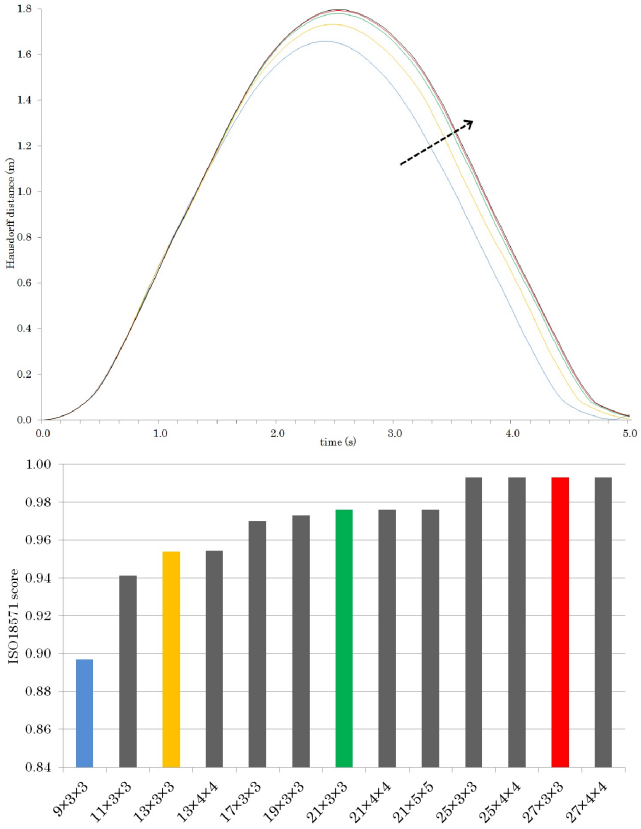

Figure 5: ***Above****: A graph showing the convergence of the NLFEM model when simulating the 3D cantilever beam scenario. The arrow represents an increase in mesh resolution, whilst the solid line represents the resolution with discretisation error minimised through the use of Z^2 error estimation.* ***Below****: The ISO18571 error metric scores of the signals in the above graph against the reference signal. Additional intermediate resolutions are included. The colours are for consistent mesh resolutions in each graph.*

does not have an effect on the accuracy, performing fewer updates per second whilst still maintaining accuracy will use fewer computational resources. Figure 6 suggests being able to use a time step of $\frac{1}{15}$s whilst maintaining accuracy.

4.2 Test PBDM Evaluations

Two PBDMs from the literature are evaluated in the framework using both the unsteady 3D cantilever beam scenario and reference PBDM from Section 4.1.2 as the appropriate components in the framework.

4.2.1 Evaluation 1

A mass-spring system (MSS) that uses a semi-implicit Euler integration method [NMK*06] is evaluated in the proposed framework. The mesh can be discretised as in Figure 7 into cubic unit cells. There are discrete masses at each corner of the cube with

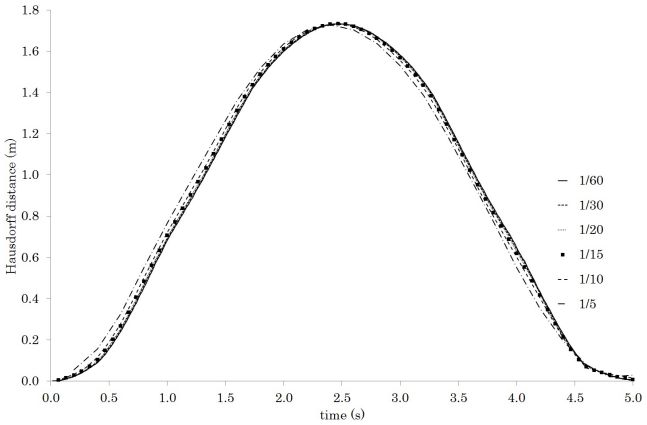

Figure 6: *The effect of varying the time step size on the measured Hausdorff distance of the displacement of the* $13\times3\times3$ *mesh for the NLFEM 3D cantilever beam scenario from its undeformed state. The solid black indicates value of convergence for the trialed timesteps.*

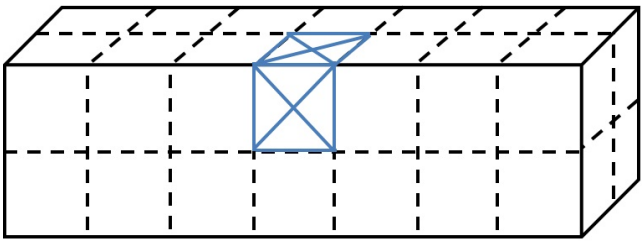

Figure 7: *A* $8\times3\times3$ *node cantilever mesh with cubes as unit cells. The springs lie across each cell's edges and face diagonals.*

springs along each edge and face diagonal. In this setup, the (linear) springs all have the stiffness coefficient k arising from the identity [KNS15]: $E = 2.5\frac{k}{a}$, where a is the length of the cubic unit cell. It is reiterated that the modelled material is homogeneous, isotropic and has a Poisson's ratio $\nu = \frac{1}{4}$.

Using a cubic unit cell for the beam can, if the discretisation is of low enough resolution, cause the beam to (incorrectly) not keep its neutral plane [KNS15]. The neutral plane is the surface between the outer surfaces (that are under compression or tension) that is not under stress. Constructing an smaller mesh (of one fewer node in each Cartesian coordinate direction) to lie inside the outer mesh, such that each node of the smaller mesh lies in the centre of each cubic unit cell of the outer mesh, overcomes this problem as springs are now more concentrated around the neutral plane of the beam. See Figure 8.

The semi-implicit Euler method is conditionally stable [Nii99] and the problem setup results in a numerically stiff system; the choice of dt depends on the point masses and the stiffness of the springs:

$$dt < C\sqrt{\frac{m}{k}},$$

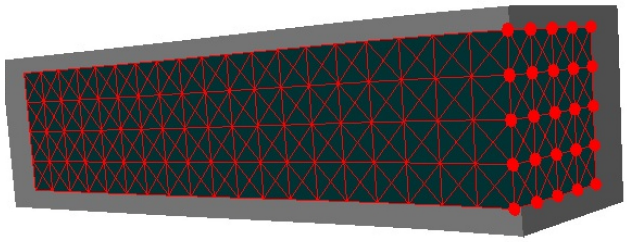

Figure 8: *The smaller mesh inside the outer mesh provides more resolution per unit length. The spheres represent nodes that are fixed (do not move).*

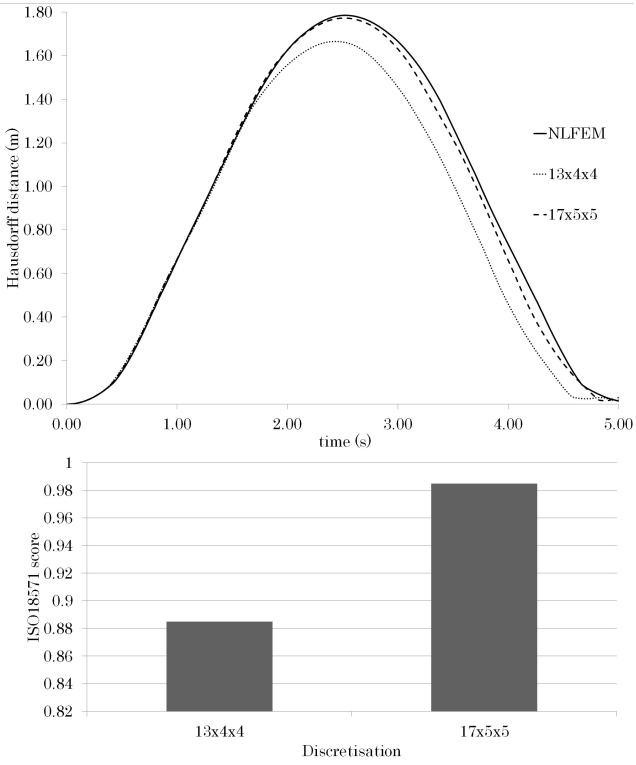

Figure 9: *The MSS is integrated using a semi-implicit Euler method.* ***Above:*** *The error measure is plotted against the reference NLFEM for different resolutions.* ***Below:*** *The ISO18571 is used to score the test (MSS) signals against the (NLFEM) reference signal.*

where C is a constant [MHTG05]. The integration is performed at fixed time intervals of $dt = \frac{1}{120}$s to accommodate this. The results for different discretisations are plotted in Figure 9 against the reference NLFEM outlined in Section 4.1.2. Higher resolution discretisations rendered the system unstable. Either a change in material properties or the adoption of another integration method is needed if either a higher resolution discretisation of the mesh or a greater time step is desired. The scores are greatest for the discretisation of

$17 \times 5 \times 5$ nodes: $e_{\text{ISO18571}} = 0.985$, which is excellent according to ISO18571.

4.2.2 Evaluation 2

A second evaluation of a test PBDM in the framework is the same MSS as in Section 4.2.1, except now the equations of motion are solved using a so-called variational implicit Euler method [MTGG11]. A discretisation of Newton's second law ($\boldsymbol{f} = m\boldsymbol{a}$) using the implicit Euler integration scheme is

$$M\frac{\boldsymbol{q}_{n+1} - 2\boldsymbol{q}_n + \boldsymbol{q}_{n-1}}{dt^2} = \boldsymbol{f}(\boldsymbol{q}_{n+1}), \tag{8}$$

where M is a (diagonal) mass matrix. Setting $\boldsymbol{x} := \boldsymbol{q}_{n+1}$ and $\boldsymbol{y} := 2\boldsymbol{q}_n - \boldsymbol{q}_{n-1}$ and rearranging (8) gives

$$M(\boldsymbol{x} - \boldsymbol{y}) = dt^2 \boldsymbol{f}(\boldsymbol{x}). \tag{9}$$

The solutions of (9) correspond to the critical points of

$$g(\boldsymbol{x}) = \frac{1}{2}(\boldsymbol{x} - \boldsymbol{y})^T M(\boldsymbol{x} - \boldsymbol{y}) + dt^2 E(\boldsymbol{x}), \tag{10}$$

where E is a potential energy field such that $\boldsymbol{f}(\boldsymbol{x}) = -\nabla E(\boldsymbol{x})$ [MTGG11] [LBOK13]. The task is then to minimise (10).

Liu *et al.* reformulate E by introducing a set of rest-length spring directions $\boldsymbol{d}_i$ (that must be found at each simulation step) and setting the system's potential energy to be the sum of each individual spring's potential energy. The task is then to optimise for both unknowns $\boldsymbol{x}$ and $\boldsymbol{d}$. A block-coordinate descent (or alternating local-global) method is employed to solve for each of $\boldsymbol{x}$ and $\boldsymbol{d}$ individually in an alternating, two-step fashion by fixing the other unknown, giving fast and stable results [LBOK13].

Since the method is implicit, it is unconditionally stable and so the time step is set to $dt = \frac{1}{60}$s, with 10 iterations of the solver per update (as set as default by the authors). The same procedure as in Section 4.2.1 is carried out, with signals given in Figure 10 together with error metric scores. The scores are excellent for a discretisation of at least $17 \times 5 \times 5$ nodes, with the optimal score being for the $29 \times 8 \times 8$ resolution. Liu *et al.* acknowledge that, for a swinging curtain scenario, just one iteration yields a "plausible" simulation but lacks particular detail and looks "a bit inflexible'. Whereas for 100 or 1000 iterations of the solver "it is difficult to tell the difference" from their reference simulation. These observations are reflected in the error metric scores for this test 3D cantilevered beam scenario on a $17 \times 5 \times 5$ mesh:

Iterations	e_{ISO18571}	Grade
1	0.910	Good
5	0.984	Excellent
10	0.984	Excellent
100	0.984	Excellent
1000	0.984	Excellent

4.3 Analysis

The smaller, inner mesh of section 4.2.1 produces an excellent grading (0.985) under the ISO18571 error metric, yet could be closer still to a maximum score of 1. This value takes into account any signal fluctuations corresponding to unavoidable factors such as numerical error. Such a high score as 0.985 is not an unrealistic expectation since the expressions are derived analytically in the literature [KNS15], matching the parameters of MSS to NLFEM. The smaller mesh is offset (see Figure 8) and, upon reconstruction to the larger, outer mesh, the fixed surface could exhibit stiffer behaviour than desired (particularly at the fixed face) that in turn influences the results. A larger resolution is unstable in this scenario but in combination with smaller timesteps a better result could be produced. This is a trade-off with performance.

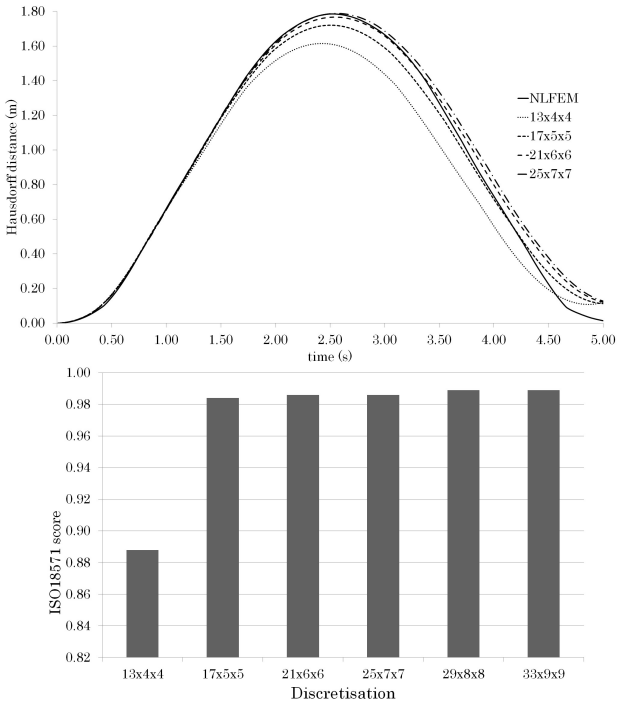

Figure 10: *The MSS is updated using a block-coordinate descent method [LBOK13]. **Above:** The Hausdorff distance for displacement is plotted against time for the test (MSS) signals at different resolutions together with the (NLFEM) reference signal. **Below:** The ISO18571 is used to score the test signals against the reference signal. Further resolutions are plotted that are absent in the above graph.*

The PBDM evaluated in Section 4.2.2 is a more approximate model to the aforementioned model; it formulates an optimisation problem that iteratively solves for a minimal result but is not guaranteed to reach a convergent solution for any number of iterations. A loss of energy is observed in Figure 10; the measured error does not return to 0 as in the reference model. This is due to the implicit integration method used in combination with the minimisation performed in the iterative solver. Nevertheless, the authors set the iteration count to 10 as it offers "the best trade-off between speed and [visual] quality". Increasing or descreasing the number of iterations generally increases or decreases the accuracy, respectively. However, to achieve real-time performance in an application, the iteration count cannot be too large. This is one trade-off that must be leveraged and the effects can be seen directly within the framework in the output scalar value. The framework allows the user to achieve their desired performance goals whilst still having a quan-

titative grasp on the accuracy of the PBDM being developed. Moreover, we can now see that in these setups, by direct comparison, the local-global iterative solver (with 10 iterations) of Section 4.2.2 is marginally less accurate than the semi-implicit integration scheme of Section 4.2.1: 0.984 < 0.985. Whereas at just one iteration, the method is less accurate than with 10 iterations: 0.910 < 0.984.

In the evaluations presented, the Hausdorff distance, popular in computer vision and shape matching, in particular, is used as the error measure at each update step. Assuming no *a priori* knowledge of the deformation, every node must be evaluated at each step. This makes for an $O(n^2)$ algorithm and is not always suitable for real-time applications. Otherwise, the user's knowledge of the specific scenario can allow for fewer nodal comparisons, reducing computational costs. Nevertheless, faster algorithms exist that can reduce the computational expenses [ZHH*17].

The scenario time of the 3D cantilever is chosen as 5 seconds to balance the allowance of a large, unsteady deformation over a substantial period of time with a fast quantitative evaluation of accuracy. That is, the time to generate quantitative results depends on how long the scenarios are set to run for. A long scenario time can hinder the user from getting rapid feedback before optimising the test PBDM. This reduces productivity in using the framework. That said, the 3D cantilever is unsteady and begins to oscillate. Any phase errors might be neglected because they have not yet established themselves before the scenario comes to a finish. This is a trade-off that is acceptable since it is assumed that a deformation that lasts longer than five seconds is not one that is typical in interactive CG applications. Nevertheless, more suitable scenario components can and should be substituted in as needed.

Using the displacements from the initial (undeformed) positions could possibly give falsely positive error metric results if the error measure is not a signed measure. In the case of two bodies deforming in opposite directions from their identical initial positions before return back to those initial positions, the error measure, being unsigned, would produce similar results. If a signed measure is used instead, the ISO18571 error metric could potentially take this antiphase into account when evaluating the phase error of the signals. As such, more suitable error measure can easily be substituted into the framework to overcome this.

5 Conclusions

A software framework has been presented that allows the objective evaluation of a proposed PBDM in CG, which can also have an applicability to fields outside of CG. The use of a single scalar value for accuracy provides an objective measure in a field that mostly uses subjective, qualitative validation through visual plausibility. The framework is composed of components that are connected in a software architecture, such that each component is a replaceable modular unit that exposes some necessary functionalities through its interfaces. A key component of the framework is the test PBDM component that allows the user to test any PBDM within the framework. The test PBDM is run through a test scenario in which the deformed positions are compared with a reference deformation over a prescribed time period. The resulting quantity scores the level of agreement between the two deformations and each score is categorised by a preset ISO standard that describes how good the agreement is.

For the evaluations conducted, the reference deformation is provided by what is established as the "gold standard": the nonlinear finite element method solving the principle of virtual displacements from elasticity. Two competing PBDMs, popular in the literature, have been evaluated within the proposed framework, giving scores that allow the direct comparison of PBDMs against one another.

The software framework allows for the fast development of test PBDMs by providing quantitative validation that is not limited by the time it takes to get qualitative validation from a field expert. The fast quantitative validation allows for an immediate, progressive iteration on the development of test PBDM's deformations to match the reference PBDM's deformation as close as possible. For example, parameter configurations specific to a PBDM can be trialed and improved in response to the resulting accuracy provided by the framework. As future work, an extension to the framework includes an automated, optimisation feedback loop that allows for the determination of best-fit parameters of a test PBDM to the reference PBDM. Moreover, the use of the framework as a diagnostic tool to help identify the sources of errors is something to be explored.

In future the framework will also contain a greater number of scenarios that will all contribute to one final output quantity for accuracy. The scenarios will be chosen such that they cover a range of different deformation types. The extension of the framework to the fields of fluid simulation and fracture simulation is possible without a fundamental change in design.

References

[BET14] BENDER J., ERLEBEN K., TRINKLE J.: Interactive simulation of rigid body dynamics in computer graphics. In *Computer Graphics Forum* (2014), vol. 33, Wiley Online Library, pp. 246–270. 2

[BFZ*13] BARBAT S., FU Y., ZHAN Z., YANG R.-J., GEHRE C.: Objective rating metric for dynamic systems. *Enhanced Safety of Vehicles, Seoul, Republic of Korea* (2013). 2, 3

[BKCW14] BENDER J., KOSCHIER D., CHARRIER P., WEBER D.: Position-based simulation of continuous materials. *Computers & Graphics 44* (2014), 1–10. 2

[BMO*14] BENDER J., MÜLLER M., OTADUY M. A., TESCHNER M., MACKLIN M.: A survey on position-based simulation methods in computer graphics. In *Computer graphics forum* (2014), vol. 33, Wiley Online Library, pp. 228–251. 2, 3

[BTC*17] BUI H. P., TOMAR S., COURTECUISSE H., COTIN S., BORDAS S.: Real-time error control for surgical simulation. *IEEE Transactions on Biomedical Engineering* (2017). 2

[CAK*14] COURTECUISSE H., ALLARD J., KERFRIDEN P., BORDAS S. P., COTIN S., DURIEZ C.: Real-time simulation of contact and cutting of heterogeneous soft-tissues. *Medical image analysis 18*, 2 (2014), 394–410. 2

[DHC*13] DUAN Y., HUANG W., CHANG H., CHEN W., TOE K. K., ZHOU J., YANG T., LIU J., TEO S. K., LIM C. W., ET AL.: Modeling and simulation of soft tissue deformation. In *International MICCAI Workshop on Computational and Clinical Challenges in Abdominal Imaging* (2013), Springer, pp. 221–230. 2

[DHC*16] DUAN Y., HUANG W., CHANG H., CHEN W., ZHOU J., TEO S. K., SU Y., CHUI C. K., CHANG S.: Volume preserved mass–spring model with novel constraints for soft tissue deformation. *IEEE journal of biomedical and health informatics 20*, 1 (2016), 268–280. 2

[DJ94] DUBUISSON M.-P., JAIN A. K.: A modified hausdorff distance for object matching. In *Pattern Recognition, 1994. Vol. 1-Conference A: Computer Vision & Image Processing., Proceedings of the 12th IAPR International Conference on* (1994), vol. 1, IEEE, pp. 566–568. 4

[FWY] FENG X., WAN W., YU X.: Analysis and comparison of methods for real-time solid deformation. *JOURNAL OF INFORMATION &COMPUTATIONAL SCIENCE 11*, 7, 2097–2107. 2

[GGW09] GEHRE C., GADES H., WERNICKE P.: Objective rating of signals using test and simulation responses. In *21st International Technical Conference on the Enhanced Safety of Vehicles, Stuttgart, Germany* (2009). 3

[Glo12] GLONDU L.: *Physically-based and Real-time Simulation of Brittle Fracture for Interactive Applications*. PhD thesis, École normale supérieure de Cachan-ENS Cachan, 2012. 2

[GM97] GIBSON S. F., MIRTICH B.: *A survey of deformable modeling in computer graphics*. Tech. rep., Technical Report, Mitsubishi Electric Research Laboratories, 1997. 2

[HDBC14] HAOUCHINE N., DEQUIDT J., BERGER M.-O., COTIN S.: Single view augmentation of 3d elastic objects. In *Mixed and Augmented Reality (ISMAR), 2014 IEEE International Symposium on* (2014), IEEE, pp. 229–236. 2

[HH06] HEIL M., HAZEL A. L.: oomph-lib–an object-oriented multi-physics finite-element library. In *Fluid-structure interaction*. Springer, 2006, pp. 19–49. 2, 5, 6

[HKR93] HUTTENLOCHER D. P., KLANDERMAN G. A., RUCKLIDGE W. J.: Comparing images using the hausdorff distance. *IEEE Transactions on pattern analysis and machine intelligence 15*, 9 (1993), 850–863. 2, 4

[KCO*03] KERDOK A. E., COTIN S. M., OTTENSMEYER M. P., GALEA A. M., HOWE R. D., DAWSON S. L.: Truth cube: Establishing physical standards for soft tissue simulation. *Medical Image Analysis 7*, 3 (2003), 283–291. 2

[KNS15] KOT M., NAGAHASHI H., SZYMCZAK P.: Elastic moduli of simple mass spring models. *The visual computer 31*, 10 (2015), 1339–1350. 7, 8

[LBOK13] LIU T., BARGTEIL A. W., O'BRIEN J. F., KAVAN L.: Fast simulation of mass-spring systems. *ACM Transactions on Graphics (TOG) 32*, 6 (2013), 214. 8

[LCAH11] LIU Y., CHEN W., ARENDT P., HUANG H.-Z.: Toward a better understanding of model validation metrics. *Journal of Mechanical Design 133*, 7 (2011), 071005. 2

[LW07] LAU K.-K., WANG Z.: Software component models. *IEEE Transactions on software engineering 33*, 10 (2007). 3

[LYD*12] LIAO X., YUAN Z., DUAN Z., SI W., CHEN S., YU S., ZHAO J.: A robust physics-based 3d soft tissue parameters estimation method for warping dynamics simulation. *AsiaSim 2012* (2012), 205–212. 2

[MADC08] MARCHAL M., ALLARD J., DURIEZ C., COTIN S.: Towards a framework for assessing deformable models in medical simulation. In *International Symposium on Biomedical Simulation* (2008), Springer, pp. 176–184. 2

[MBP14] MUGUERCIA L., BOSCH C., PATOW G.: Fracture modeling in computer graphics. *Computers & Graphics 45* (2014), 86–100. 2

[MHTG05] MÜLLER M., HEIDELBERGER B., TESCHNER M., GROSS M.: Meshless deformations based on shape matching. *ACM transactions on graphics (TOG) 24*, 3 (2005), 471–478. 1, 7

[MM07] MOORE P., MOLLOY D.: A survey of computer-based deformable models. In *Machine Vision and Image Processing Conference, 2007. IMVIP 2007. International* (2007), IEEE, pp. 55–66. 2

[Mon92] MONAGHAN J. J.: Smoothed particle hydrodynamics. *Annual review of astronomy and astrophysics 30*, 1 (1992), 543–574. 2

[MTGG11] MARTIN S., THOMASZEWSKI B., GRINSPUN E., GROSS M.: Example-based elastic materials. In *ACM Transactions on Graphics (TOG)* (2011), vol. 30, ACM, p. 72. 8

[Nii99] NIIRANEN J.: Fast and accurate symmetric euler algorithm for electromechanical simulations. *Proc. Electrimacs' 99, Lisboa, Portugal, 9* (1999). 7

[NLB*07] NUSSECK M., LAGARDE J., BARDY B., FLEMING R., BÜLTHOFF H. H.: Perception and prediction of simple object interactions. In *Proceedings of the 4th symposium on Applied perception in graphics and visualization* (2007), ACM, pp. 27–34. 1, 2

[NMK*06] NEALEN A., MÜLLER M., KEISER R., BOXERMAN E., CARLSON M.: Physically based deformable models in computer graphics. In *Computer graphics forum* (2006), vol. 25, Wiley Online Library, pp. 809–836. 1, 2, 6

[OT02] OBERKAMPF W. L., TRUCANO T. G.: Verification and validation in computational fluid dynamics. *Progress in Aerospace Sciences 38*, 3 (2002), 209–272. 1

[PIS*13] PATETE P., IACONO M. I., SPADEA M. F., TRECATE G., VERGNAGHI D., MAINARDI L. T., BARONI G.: A multi-tissue mass-spring model for computer assisted breast surgery. *Medical engineering & physics 35*, 1 (2013), 47–53. 2

[PP05] PILONE D., PITMAN N.: *UML 2.0 in a Nutshell*. " O'Reilly Media, Inc.", 2005. 3

[SKH*10] SARIN H., KOKKOLARAS M., HULBERT G., PAPALAMBROS P., BARBAT S., YANG R.-J.: Comparing time histories for validation of simulation models: error measures and metrics. *Journal of dynamic systems, measurement, and control 132*, 6 (2010), 061401. 2, 3

[SM15] SANKARARAMAN S., MAHADEVAN S.: Integration of model verification, validation, and calibration for uncertainty quantification in engineering systems. *Reliability Engineering & System Safety 138* (2015), 194–209. 2

[SPQ*17] SUI Y., PAN J. J., QIN H., LIU H., LU Y.: Real-time simulation of soft tissue deformation and electrocautery procedures in laparoscopic rectal cancer radical surgery. *The International Journal of Medical Robotics and Computer Assisted Surgery* (2017). 3

[SSP07] SOLENTHALER B., SCHLÄFLI J., PAJAROLA R.: A unified particle model for fluid–solid interactions. *Computer Animation and Virtual Worlds 18*, 1 (2007), 69–82. 2

[SWU*15] SCHOELL S. L., WEAVER A. A., URBAN J. E., JONES D. A., STITZEL J. D., HWANG E., REED M. P., RUPP J. D.: Development and validation of an older occupant finite element model of a mid-sized male for investigation of age-related injury risk. *Stapp car crash journal 59* (2015), 359. 3

[VDSG15] VAVALLE N. A., DAVIS M. L., STITZEL J. D., GAYZIK F. S.: Quantitative validation of a human body finite element model using rigid body impacts. *Annals of biomedical engineering 43*, 9 (2015), 2163–2174. 3

[W*12] WESTWOOD J., ET AL.: Towards a generic framework for evaluation and comparison of soft tissue modeling. *Medicine Meets Virtual Reality 19: NextMed 173* (2012), 116. 2

[Wei17] WEI Z.: Analysis, modeling and cae validation of vehicle crashes using advanced signal processing tools. 3

[YYT*15] YU S., YUAN Z., TONG Q., LIAO X., BAI Y.: Improved compressed sensing based 3d soft tissue surface reconstruction. In *Pacific Rim Conference on Multimedia* (2015), Springer, pp. 549–558. 2

[ZHH*17] ZHANG D., HE F., HAN S., ZOU L., WU Y., CHEN Y.: An efficient approach to directly compute the exact hausdorff distance for 3d point sets. *Integrated Computer-Aided Engineering 24*, 3 (2017), 261–277. 9

[ZZ92] ZIENKIEWICZ O. C., ZHU J. Z.: The superconvergent patch recovery and a posteriori error estimates. part 1: The recovery technique. *International Journal for Numerical Methods in Engineering 33*, 7 (1992), 1331–1364. 2, 6

[ZZGC17] ZHANG J., ZHONG Y., GU C., COLOE P.: Deformable models for surgical simulation: a survey. *IEEE Reviews in Biomedical Engineering* (2017). 1, 2