Evaluating Hex-mesh Quality Metrics via Correlation Analysis

Xifeng Gao\(^1\) Jin Huang\(^2\) Kaoji Xu\(^3\) Zherong Pan\(^4\) Zhigang Deng\(^3\) Guoning Chen\(^3\)

\(^1\) New York University, \(^2\) Zhejiang University
\(^3\) University of Houston, \(^4\) University of North Carolina at Chapel Hill

Abstract

Hexahedral (hex-) meshes are important for solving partial differential equations (PDEs) in applications of scientific computing and mechanical engineering. Many methods have been proposed aiming to generate hex-meshes with high scaled Jacobians. While it is well established that a hex-mesh should be inversion-free (i.e., have a positive Jacobian measured at every corner of its hexahedron), it is not well-studied that whether the scaled Jacobian is the most effective indicator of the quality of simulations performed on inversion-free hex-meshes given the existing dozens of quality metrics for hex-meshes. Due to the challenge of precisely defining the relations among metrics, studying the correlations among different quality metrics and their correlations with the stability and accuracy of the simulations is a first and effective approach to address the above question. In this work, we propose a correlation analysis framework to systematically study these correlations. Specifically, given a large hex-mesh dataset, we classify the existing quality metrics into groups based on their correlations, which characterizes their similarity in measuring the quality of hex-elements. In addition, we rank the individual metrics based on their correlations with the accuracy and stability metrics for simulations that solve a number of elliptic PDE problems. Our preliminary experiments suggest that metrics that assess the conditioning of the elements are more correlated to the quality of solving elliptic PDEs than the others. Furthermore, an inversion-free hex-mesh with higher average quality (measured by any quality metrics) usually leads to a more accurate and stable computation of elliptic PDEs. To support our correlation study and address the lack of a publicly available large hex-mesh dataset with sufficiently varying quality metric values, we also propose a two-level perturbation strategy to generate the desired dataset from a small number of meshes to exclude the influences of element numbers, vertex connectivity, and volume sizes to our study.

Categories and Subject Descriptors (according to ACM CCS):
I.3.3 [Computer Graphics]: Computational Geometry and Object Modeling—Curve, surface, solid, and object representations

1. Introduction

Hexahedral (hex-) meshes are commonly employed for solving volumetric partial differential equations (PDEs) due to their desirable numerical properties [BPM\(^\ast\)95], such as their natural support of tensor products [ZBG\(^\ast\)07, LZW\(^\ast\)14], which usually leads to efficient computations [Tanu\(^\ast\)04]. Aiming at producing high-quality hex-meshes, a number of techniques have been proposed [SJ\(^\ast\)08, SLSK\(^\ast\)04, ZZM\(^\ast\)07, Mar\(^\ast\)09, ZLX\(^\ast\)13, NRP\(^\ast\)11, HTWB\(^\ast\)11, LLL\(^\ast\)12, JHW\(^\ast\)14, LKB\(^\ast\)16, GSS\(^\ast\)11, WLL\(^\ast\)12, LV\(^\ast\)13, LLW\(^\ast\)13, HIS\(^\ast\)14, GDC\(^\ast\)15, FXB\(^\ast\)16, GMD\(^\ast\)16, LMP\(^\ast\)16, FBL\(^\ast\)16]. While dozens of metrics for evaluating the quality of a hex-mesh have been proposed [Knu\(^\ast\)00a, Knu\(^\ast\)00b, SEK\(^\ast\)07, Joe\(^\ast\)08, MJN\(^\ast\)11], the scaled Jacobian is predominantly used in the above mentioned hex-meshing techniques to evaluate the quality of the produced hex-meshes. It is generally acknowledged that a hex-mesh with no inverted elements (i.e., with a positive minimum scaled Jacobian), is a hard requirement to conduct PDE-based simulations [Cia\(^\ast\)02, SJ\(^\ast\)08, PTS\(^\ast\)08] on it. Once this minimum requirement is satisfied, however, is the scaled Jacobian still the most effective quality indicator for a hex-mesh? To the best of our knowledge, such an important query has not been sufficiently addressed, which requires the understanding of the similarity among existing metrics and the study of their capability of implying the quality of simulations performed on the given hex-mesh.

Considering the lack of a theoretical formulation of the relations of qualities between meshes and simulations, we argue that the first and effective step to understand their relations is to inspect their covariant behaviors, which we wish will enlighten the development of the missing theory. Consequently, we propose to statistically investigate the effectiveness of the available quality metrics \(\dagger\) for a hex-mesh via correlation analysis (e.g., the Pearson product-moment correlation coefficient (PPMCC) [Sti\(^\ast\)89]), which is the main contribution of this work. Specifically, we perform the correlation analysis based evaluations using two strategies: (1) evaluate the correlations \(\dagger\) We employ all the quality metrics for a hexahedral element provided in [SEK\(^\ast\)07] and consider they are inclusive. However, any missing quality metrics can be incorporated into our framework seamlessly.

© 2017 The Author(s)
Computer Graphics Forum © 2017 The Eurographics Association and John Wiley & Sons Ltd. Published by John Wiley & Sons Ltd.
tions between any pair of metrics without considering downstream applications (Section 4.1) and (2) investigate the impact of a metric to the quality (measured by eigenvalues of the system) of the elliptic PDE-based applications (Section 4.2). In contrast to previous works that evaluate a few metrics on a single element (e.g., a tetrahedral or a hexahedral) [Knu00a, Knu00b, She02, Joe08], our study focuses on understanding the influences of a large set of quality metrics of the entire mesh to the conditioning of FEM and the accuracy of interpolation functions, which distinguishes our work from existing quality metric evaluation work in the literature.

To perform the proposed statistical analysis, a large hex-mesh dataset with well-distributed metric values is needed. Ideally, hex-meshes produced by hex-meshing techniques applied in industry should be used to create the test dataset. However, none of the existing technique can robustly generate a large number of hex-meshes with sufficiently varying metric values, which will introduce strong bias to the correlation analysis. To address this, we propose a two-level perturbation strategy to generate the desired dataset (Section 3), which is our second contribution. This perturbation strategy excludes the influences of element numbers, vertex connectivity and volume sizes to the quality of simulations performed on hex-meshes, enabling us to concentrate on the evaluation of shape metrics.

From our evaluation, we discover that metrics that are relevant to the conditioning of the metric tensor (i.e. $J^2$ where $J$ is the Jacobian matrix [Knu00b]) of the elements are more correlated to the quality of solving elliptic PDEs than the others. Furthermore, an inversion-free hex-mesh (i.e. with positive minimum scaled Jacobian) with higher average quality (measured by any metrics) usually leads to a more accurate and stable computation of elliptic PDEs. Both discoveries are not commonly known by the meshing community, which we consider as another important contribution of this work. Note that, our conclusions are derived from the analysis of artificially generated datasets, which might not be valid for real world meshes. Also, because of the employment of the eigenvalue-based metrics for simulation quality and the practically-less-relevant L2 norm error measurement, the discovery relations of the metrics with the simulation quality may not be practical. All these need further validations.

We believe the proposed framework – quantitatively evaluating a set of metrics via correlation analysis, is general enough and can be applied to study the quality metrics of a mesh represented in other element types (i.e. polygons and polyhedra). The source code of the framework, including the PDEs solving on hex-meshes and the perturbation-based dataset generation, will be released to foster future studies in this direction.

2. Related Work

Hex-meshing: Generating an all-hex mesh that consists of hexahedral elements has been studied for decades. In practice, users still greatly rely on semi-automatic/interactive hex-meshing techniques [San16, SJ08, GMD'16, LMPS16] to produce coarse structured hex-meshes or employ grid/octree-based methods [SLSK04, ZZM07, Mar09, ZLX13] to create highly unstructured ones. Recently, researchers have made great progresses to automatically generate feature-aligned hex-meshes with the aid of volumetric parameterization, such as polycube based methods [GSZ21, WLL’12, LVS’13, LLWQ13, HJS’14, FXBH16, FBL16] and frame field based approaches [NRP11, HTWB11, LXX’12, JHW’14, LKB16]. All these techniques are designed to generate hex-meshes consisting of hexahedra with shapes that are as close as possible to canonical cubes. However, since the shape of the boundary of an input object is arbitrary, the current hex-meshing techniques cannot guarantee the generated hex-mesh is inversion-free. Therefore, a set of hex-mesh optimization techniques have been proposed to improve the values of quality measurements, such as the scaled Jacobian [Knu00c, Knu03, BDK’03, RGRS14, LSVT15, GC16] and the condition number [Knu03, BDK’03]. However, once all the hexahedra of a hex-mesh are valid, will continually improving the scaled Jacobian of a hex-mesh result in more accurate and more stable simulations performed on it? Why the scaled Jacobian is typically used in contrast to the other metrics? Up to now, such important questions remain to be answered.

Table 1: Statistics of hex-mesh quality metrics in [SEK*07] (pp.82–100). Range columns with and without * respectively represent the actual and user-specified bounds.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Abbr.</th>
<th>Range</th>
<th>Range*</th>
<th>Trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>diagonal</td>
<td>D</td>
<td>[0, 1]</td>
<td>[0, 1]</td>
<td>↑</td>
</tr>
<tr>
<td>dimension</td>
<td>DM</td>
<td>[0, +∞]</td>
<td>[0, +∞]</td>
<td>↑</td>
</tr>
<tr>
<td>distortion</td>
<td>DIS</td>
<td>[0, +∞]</td>
<td>[0, 1] 8</td>
<td>↑</td>
</tr>
<tr>
<td>edge ratio</td>
<td>ER</td>
<td>[1, +∞]</td>
<td>[1, +∞]</td>
<td>↓</td>
</tr>
<tr>
<td>Jacobian</td>
<td>J</td>
<td>[1, +∞]</td>
<td>[0, +∞] 8</td>
<td>↑</td>
</tr>
<tr>
<td>maximum edge ratio</td>
<td>MER</td>
<td>[1, +∞]</td>
<td>[1, +∞] 8</td>
<td>↓</td>
</tr>
<tr>
<td>aspect Frobenius</td>
<td>AF</td>
<td>[1, +∞]</td>
<td>[1, +∞] 8</td>
<td>↓</td>
</tr>
<tr>
<td>mean aspect Frobenius</td>
<td>MAF</td>
<td>[1, +∞]</td>
<td>[1, +∞] 8</td>
<td>↓</td>
</tr>
<tr>
<td>Oddity</td>
<td>O</td>
<td>[0, +∞]</td>
<td>[0, +∞] 8</td>
<td>↓</td>
</tr>
<tr>
<td>relative size squared</td>
<td>RSS</td>
<td>[0, 1]</td>
<td>[0, 1]</td>
<td>↑</td>
</tr>
<tr>
<td>scaled Jacobian</td>
<td>S. J</td>
<td>[1, 1]</td>
<td>[0, 1]</td>
<td>↑</td>
</tr>
<tr>
<td>shape</td>
<td>S</td>
<td>[0, 1]</td>
<td>[0, 1]</td>
<td>↑</td>
</tr>
<tr>
<td>shape size</td>
<td>SS</td>
<td>[0, 1]</td>
<td>[0, 1]</td>
<td>↑</td>
</tr>
<tr>
<td>shear</td>
<td>SE</td>
<td>[0, 1]</td>
<td>[0, 1]</td>
<td>↑</td>
</tr>
<tr>
<td>shear size</td>
<td>SES</td>
<td>[0, 1]</td>
<td>[0, 1]</td>
<td>↑</td>
</tr>
<tr>
<td>skew</td>
<td>SK</td>
<td>[0, 1]</td>
<td>[0, 1]</td>
<td>↓</td>
</tr>
<tr>
<td>stretch</td>
<td>ST</td>
<td>[0, 1]</td>
<td>[0, 1]</td>
<td>↑</td>
</tr>
<tr>
<td>taper</td>
<td>T</td>
<td>[0, +∞]</td>
<td>[0, +∞] 8</td>
<td>↓</td>
</tr>
<tr>
<td>volume</td>
<td>V</td>
<td>[0, +∞]</td>
<td>[0, +∞] 8</td>
<td>–</td>
</tr>
</tbody>
</table>

Quality Metrics: Given a hexahedron with eight corners, its quality can be measured through various quantities, e.g., the determinant of the Jacobian matrix constructed at every corner, the ratios between the determinants, angles and so on. Dozens of metrics are derived from these quantities [Knu00a,Knu00b,SEK*07,Joe08,MNI11]. For each measurement, $X$, there are generally two corresponding metrics to evaluate the quality of a hex-mesh, i.e., an extreme metric to locate the element of the hex-mesh with the worst quality, denoted as $X_{\text{Min}}/X_{\text{Max}}$, and an average metric to estimate the overall quality of all hexahedra in the hex-mesh, denoted as $X_{\text{Avg}}$. We have summarized these metrics in Table 1. The Metric, Abbr., and Range columns of Table 1 stand for the names, abbreviations, and value ranges of the metrics, respectively. We consider the left value of the
range listed in the Range column of Table 1 for a metric as its lower bound, while the right value as its upper bound only when it is not $+\infty$. Note that, the lower bounds in the Range column of Table 1 are constrained to satisfy the inversion-free requirement of the to-be-generated hex-meshes. For the Trend column, $\uparrow$ means a higher value of the corresponding metric is preferred, while $\downarrow$ means the opposite. While paper [SEK'07] describes the calculations of most of the listed metrics, in the supplemental document, we provide the computations of several important ones and the Dimension metric that is not straightforward to understand. Given these metrics, users often have to balance the requirements of specific applications. For example, a highly inhomogeneous anisotropic problem may want some elements in “ugly shapes”. The efficiency and robustness of some direct solvers (e.g. Cholesky decomposition) are less dependent on the global structure and spectral features (eigenvalues) than some iterative solvers (especially Krylov space methods [Gut07] and multi-grid methods [Fal06]). In this paper, for the first time, we introduce a general framework to conduct a comprehensive study to understand the relations among these metrics, the characteristics of the elements they measure, and their effectiveness in measuring the quality of a hex-mesh.

**Hex-mesh Evaluation:** During the past decades, intensive studies have been conducted on the comparisons of tet-meshes and hex-meshes with respect to their numerical properties, computational efficiency, and rate of convergences [CK92, BPM'95, RS06, BTPB07, TEC10, TEC11, Cha13], respectively. However, to date, there exist few studies on the effectiveness of the various metrics for hex-meshes. Knupp [Knu00b] evaluated their optimization framework for hex-meshes using a number of objective functions based on the local Jacobian matrix. Muller et al. [MHKSZ01] investigated the effect of the average scaled Jacobian and the number of hex-elements on the accuracy of a linear elasticity based simulation. They concluded that both of the number of hex-mesh elements and the average scaled Jacobian have positive impacts on the convergence and the accuracy of the simulations, which is independently verified by our experiments (Figure 1 and 6). In addition, our study verifies this positive impact of average scaled Jacobian in two other elliptic PDE problems, i.e., Poisson’s equation solving and Stokes equation solving. By performing FEA on electromagnetics, Motooka et al. [MN11] compared four quality metrics evaluated based on the orthogonality, facet planarity, diagonal length ratio, and volume ratio of each hex element in the hex-mesh, respectively. They concluded that the diagonal length ratio affected the convergence of the solver the most, which shares some similarity to our discoveries made from this study. However, our discoveries are more general and identify that metrics relevant to the conditioning of the element are more effective. Owen et al. [OS14] compare performances of elastic plastic simulations performed on hex-meshes generated by the grid-based and the traditional pave-and-sweep based approaches. In contrast to previous works that evaluate quality metrics either using a small set of data or lacking verifications in terms of subsequent applications, our work generates a large hex-mesh dataset and studies the correlations among the quality metrics and between the mesh quality metrics and simulation quality via three important applications, including the linear elasticity problem [NMK'06], Poisson’s equation solving [KBH06] and Stokes equation solving [BB10, ZYZZ15], respectively.

3. **Dataset Preparation**

To ensure a fair and reliable conclusion from our metric evaluation, we enforce the following requirements to a dataset:

1. All hex-meshes in the same dataset should have the same number of elements and volume. This is because the values of most of the shape metrics we evaluate will be strongly affected by how the mesh is discretized [MHKSZ01] or how large the mesh is, which will affect the accuracy and stability of the simulations. For example, Figure 1 shows that the errors and instability of four applications (detailed in Section 4.2.2) performed on a cube model are monotonically reduced with the increase of the number of elements in the cube hex-meshes.

![Figure 1](image-url)

\* Figure 1: Both the accuracy and the stability of the elliptic PDE problems are improving as the element number of the cube hex-mesh increases. X-axis is the element number of a hex-mesh. Y-axis is either $\log_{10}(\lambda_{\text{min}})$ for the accuracy of a solution, $\log(\lambda_{\text{max}})$ for the stability of a solution, or the true value of the “Analytical Solution”.

2. We only take “valid” meshes into account, i.e. the minimum scaled Jacobian (i.e. S.JAC.Min) of a valid hex-mesh is positive. This property is required simply because that an invalid hex-mesh cannot be used for the solving of the subsequent elliptic PDEs [Ciu02, PTS'08] and our study focuses on whether S.JAC.Min is still an effective metric for a valid hex-mesh.

3. The dataset of the hex-meshes should cover a value range for each quality metric as wide as possible to ensure sufficient variation for the subsequent correlation analysis. However, there doesn’t exist a hex-meshing technique to build a dataset with the desired metric distributions while satisfying the above two criteria. For instance, a dataset (refer to the supplementary) generated from MeshGem [Mes15], which is octree-based – one of the most robust and efficient methods, lacks enough variance on many metrics (e.g., the difference between the maximal and minimal values of S.JAC.Min is only 0.074) and the produced hex-meshes also have different numbers of elements.

Inspired by previous evaluation works [MHKSZ01, MN11], to generate a dataset satisfying the above criteria, we propose a hierarchical, two-level geometric-perturbation algorithm.

Note that a hex-mesh is associated with a unique vector $\vec{U}$ containing all its metric values. Each entry value of $\vec{U}$ falls in the range determined by the lower and upper bounds of the corresponding metric in Table 1. The basic idea of our data generation strategy is to start with an input hex-mesh produced by a state-of-the-art
technique (e.g. the polycube mapping or the frame field based approach) or manually designed that has a best possible quality, $\hat{U}_b$. The entries of $\hat{U}_b$ provide an upper bound—if it corresponds to a minimum metric, or a lower bound—if it corresponds to a maximum metric for the quality of the to-be-generated hex-meshes. Our goal is to generate hex-meshes with metric values falling in the ranges determined by $\hat{U}_b$ while having sufficient variance. To achieve that, we first generate a set of hex-meshes $H_{\text{min}}$ by explicitly controlling their values of those minimum metrics with finite bounds, then from the meshes in $H_{\text{min}}$ we generate a large mesh dataset $H_{\text{ave}}$ by explicitly controlling their average metric values. Note that for metrics without finite bounds (e.g. those maximum metrics), we do not explicitly control their distribution. In the following, we provide more details on our data generation strategy.

**Generating dataset $H_{\text{min}}$:** Given an input valid hex-mesh $H_0$ with metric vector $\hat{U}_b$, we evenly subdivide each of its minimum metric value ranges into $\alpha$ sub-ranges. $\alpha$ is used to control the amount of variation in $H_{\text{min}}$. The larger the value of $\alpha$, the more variation is in $H_{\text{min}}$, but more computation is needed. In our experiment, we set $\alpha = 8$ to achieve a good tradeoff between sufficient variation and computation cost. The to-be-generated dataset $H_{\text{min}}$ is composed of hex-meshes with varying index vectors, $I_s$. The entries of each $I_s$ correspond to the sub-range ids of the minimum metrics that the respective metric value of the associated hex-mesh fall in. To generate a hex-mesh $H'$ from $H_0$, we perturb the interior vertices of $H_0$ one-by-one in a random order by adding a vector $\vec{r}$ to the vertex $(|\vec{r}| < s$ and $s$ is a percentage of the average edge length of $H_0$, e.g., 0.1%). After perturbing a vertex, we compute an integer vector $\vec{f}$ of $H'$. If $\vec{f}$ is different from all the $I_s$ corresponding to the current hex-meshes in $H_{\text{min}}$, we add $H'$ to $H_{\text{min}}$. Otherwise, for a metric $m_j$ and its sub-range that currently does not have a mesh in $H_{\text{min}}$ corresponding to, if the metric value of $m_j$ of $H'$ is larger than the upper bound of this sub-range, we set $s \leftarrow 2s$ and perform another perturbation. Similarly, if the metric value of $m_j$ of $H'$ is smaller than the lower bound, we set $s \leftarrow s/2$ and perform the next perturbation. This trial is performed for 5 times and returns null if no new mesh can be found.

Ideally, the above perturbation process should generate $\alpha \times N_{\text{min}}$ meshes with each sub-range covered by a hex-mesh. $N_{\text{min}} = 11$ is the number of minimum metrics in Figure 2. Even with the above adjustment strategy of the amount of perturbation, a few sub-ranges may be still not covered by any hex-meshes. Therefore, the number of the generated hex-meshes is typically smaller than $\alpha \times N_{\text{min}}$. Figure 2(top) visualizes the distributions of 67 hex-meshes ($H_{\text{min}}$) of the crank model using parallel coordinates [Ins85]. From Figure 2, we see that the generated hex-meshes have well distributed values across the ranges of those minimum metrics but approximately constant values of the average metrics, as only a few hex elements are altered. Next, we generate the complete dataset from these meshes in $H_{\text{min}}$.

**Generating dataset $H_{\text{ave}}$:** Similarly, for each hex-mesh in $H_{\text{min}}$, we divide each of the average metric value range into $\beta$ sub-ranges and use the same perturbation strategy described earlier to the interior vertices. The number of the perturbed average metrics is $N_{\text{ave}} = 10$. Because of the same reason as in the preceding step, the actual number of the obtained valid meshes is smaller than the ideal number $\beta \times N_{\text{ave}}$. The $H_{\text{ave}}$ row of Table 2 lists the numbers of the generated hex-meshes for six models after performing this step. Given the limited space and time, we randomly select less than 1000 meshes from $H_{\text{ave}}$, from which we further remove the outliers using the parallel coordinate tool [Ins85]. The selected meshes, $F_{\text{ave}}$, form the dataset for the corresponding model. The $F_{\text{ave}}$ row in Table 2 shows the numbers of hex-meshes in six representative datasets for the evaluation. Figure 2(bottom) visualizes the distribution of 818 hex-meshes ($F_{\text{ave}}$) of the crank model. From Figure 2(bottom), we see that the generated hex-meshes have reasonably well distributed values across the ranges of the average metrics that we control.
Table 2: Statistics of six datasets. #H represents the number of elements in a mesh. $H_{\text{Max}}$, $H_{\text{Avg}}$ and $F_{\text{Avg}}$ indicate the number of hex-meshes generated in different stages of the data generation, respectively. Timing shows the time for data generation, while Timing$^*$ is the total time for simulations.

<table>
<thead>
<tr>
<th>Datasets</th>
<th>Bone</th>
<th>Bust</th>
<th>Elephant</th>
<th>Hanger</th>
<th>Bunny</th>
<th>Rockerarm</th>
</tr>
</thead>
<tbody>
<tr>
<td>#H</td>
<td>3396</td>
<td>5398</td>
<td>8730</td>
<td>4539</td>
<td>4552</td>
<td>5993</td>
</tr>
<tr>
<td>$H_{\text{Max}}$</td>
<td>64</td>
<td>57</td>
<td>49</td>
<td>62</td>
<td>60</td>
<td>59</td>
</tr>
<tr>
<td>$H_{\text{Avg}}$</td>
<td>3724</td>
<td>2752</td>
<td>3795</td>
<td>3654</td>
<td>3903</td>
<td>4554</td>
</tr>
<tr>
<td>$F_{\text{Avg}}$</td>
<td>525</td>
<td>607</td>
<td>643</td>
<td>719</td>
<td>735</td>
<td>565</td>
</tr>
<tr>
<td>Timing</td>
<td>24h</td>
<td>40h</td>
<td>4h</td>
<td>51h</td>
<td>34h</td>
<td>61h</td>
</tr>
<tr>
<td>Timing$^*$</td>
<td>8.3h</td>
<td>10.5h</td>
<td>4.13h</td>
<td>10.5h</td>
<td>8.5h</td>
<td>12.5h</td>
</tr>
</tbody>
</table>

Final datasets: We generate a dataset for each of the 22 objects with various geometric and topological complexities (Figure 4(b)) using the above described strategy. Table 2 provides the statistics for six datasets. Statistics for all the datasets can be found in the supplemental material. The data generation and the simulations (Section 4.2) were computed on two PCs – PC1 has an Intel Xeon (E5-1620 @ 3.70GHz) CPU and 48GB memory, while PC2 has an Intel Xeon (E5-2640 @ 2.60GHz) CPU and 128GB memory. For each dataset, its generation and simulations took at most 3 days in total. The total size of the generated datasets is about 500GB \(^\dagger\).
4.1. Application Independent Study

Our first evaluation computes the correlations among different metrics that are hard to derive due to the lack of analytical solutions, based on which we classify metrics into different groups to study their common properties.

For each of the 22 datasets, we compute its correlation matrix of the metrics listed in Table 1. According to their correlation matrices, we classify all the 22 datasets into a few groups, as shown in Figure 4(a). We used Ward’s clustering technique [ML14] as detailed in the next paragraph. From this clustering result, we can hardly observe any common characteristics for the objects within each group, as also evidenced by the deviation matrix with respect to their average matrix (see the supplementary document). Therefore, to minimize the potential bias caused by a specific model to the relationships among metrics, we use the average correlation matrix of all the correlation matrices for later evaluations. Figure 4(a) visualizes this matrix. Each entry of the matrix is represented by a colored dot with red and blue indicating positive and negative correlations, respectively. The darker the color is, the stronger the correlation is. Also, the sizes of dots indicate the absolute correlation coefficient values. The larger the dots, the stronger the correlations that they represent.

While the similarity of any pair of metrics is measured by the correlation coefficient, it is not appropriate to group these two metrics together purely because their correlation is strong, as their respective correlations to other metrics will in turn affect their correlation. For instance, consider metrics $m_1$, $m_2$ and $m_3$ and their pairwise correlation coefficients as $r_{m_1, m_2} = 0.9$, $r_{m_1, m_3} = 0.8$ and $r_{m_2, m_3} = 0.46$. It is apparent that the similarity measurement between metrics $m_1$ and $m_2$ cannot be simply described by $r_{m_1, m_2}$ due to the influence of $m_3$. In our case, for each quality metric, we consider its corresponding row or column in the correlation matrix (Figure 6) with absolute entry values as its feature vector. Given two metrics and their respective feature vectors, $\vec{i}$ and $\vec{j}$, if they have high similarity which is measured by Eq.(2) (refer to [KR09] for details), then these two metrics should also have similar evaluation capacity to the hex-meshes.

$$d(\vec{i}, \vec{j}) = \left| \frac{|i|}{|i| + |j|} \right| \sqrt{i^2 + j^2},$$

where $|.|$ and $\|.|\|$ denote absolute value and norm, respectively.

We then perform the Ward’s hierarchical agglomerative clustering [ML14] to organize the metrics into a binary tree based on this dissimilarity measure. Figure 5 shows the obtained tree structure describing the hierarchical clustering of the metrics. We classify these metrics into 8 groups (highlighted by the red rectangles in Figure 5), with the distance between any two metrics in the same cluster less than 1. Note that the distance among any pair of metrics is shown in the Height (or Y) axis.

Based on this clustering, we decompose the correlation matrix shown in Figure 4(a) into three sub-matrices, which show the correlations among Min./Max. metrics (Figure 6(a)), among Avg. metrics (Figure 6(b)), and between the Min./Max. metrics and their corresponding Avg. metrics (Figure 6(c)).

From Figures 6 and 5, we have the following observations:

1. All average metrics are classified into the same cluster.
2. There are strong correlations with each other (Figure 6(b)). This shows that when measuring the...
overall quality of a hex-mesh, all shape metrics defined for a hexahedral element have similar capability.

(2) Minimum/maximum metrics that measure the extreme quality of an element have different correlations (Figure 6(a)). Based on different geometric characteristics of the elements that those metrics focus on, they can be classified into different groups (Figure 5). For instance, metrics that measure the shape, size and shearing based on element volumes belong to the same clusters (e.g. DM. Max. and VOL. Max. in cluster 1, SPS. Min. and SES. Min. in cluster 7 and RSS. Min. and VOL. Min. in cluster 8, respectively). ODDY Max., AF. Max. and MAF. Max. belong to cluster 5 as they measure the deviations of the metric tensor [Knu00b] from the identity matrix either directly or indirectly. Metrics that calculate the ratios of edge/diagonal lengths of principal axes are also highly correlated (e.g. ER. Max. and ST. Min. in cluster 3, DIAG. Min., MER. Max., SKEW. Max., and TAP. Max. in cluster 5, respectively), while those that measure angle distortions are correlated (e.g. JAC. Min., DSTT. Min., SE. Min., JAC. Avg., and SP. Min. in cluster 6).

(3) Most Min/Max metrics have positive correlations with their corresponding Avg. metrics, as shown by the diagonal entries in Figure 6(c), except for DM. Max. that has negative correlation with its average metric. Metrics ER. Max., ST. Min., DIAG. Min., MER. Max., SKEW. Max., TAP. Max., MAF. Max., DM. Max., VOL. Max. all have strong correlations with their respective average metrics AND all other average metrics as well, while the rest of the metrics has small correlations even with their corresponding average metrics. In addition, metric RSS. Min. has very small correlations with other metrics, except for VOL. Min. (Figure 6(a)). This is because RSS (i.e. relative size squared) measures the ratio of the hex volume to the average volume, thus RSS. Min. is proportional to VOL. Min. Both RSS. Min. and VOL. Min. have almost no correlation with all average metrics (Figure 6(c)).

Conclusions of our application independent study: With this clustering result and the above observations, researchers focusing on hex-mesh generation techniques can use a smaller number of metrics to evaluate the quality of the generated hex-meshes, i.e., selecting one representative metric from each cluster based on the characteristics of the mesh that they are interested in. For example, if one cares about angle distortion of the mesh, either SE. Min. or S.JAC. Min. is needed. If the deviations of the metric tensor from the identity matrix is of interest, one metric from cluster 5 is sufficient. This study characterizes the common properties of certain metrics, which is shown very important to understand their effectiveness in the specific applications (Section 4.2). Furthermore, this study provides additional discovery (i.e. the above observation (3)) that has not been discovered before.

4.2. Application Dependent Study

After studying the correlations among metrics, we further investigate their influences on three well-studied applications, i.e., linear elasticity problem, the Poisson’s equation and Stokes equation solving. Other applications can be analyzed in a similar way. In the following subsections, we first briefly introduce these applications, then detail our evaluations considering these applications.

4.2.1. Elliptic PDEs

The Linear Elasticity models an elastically deformable body under infinitesimal displacement or traction. According to Bathe [Bat95], isotropic linear elastic energy is a quadratic energy of the displacement field $v$, taking the following form:

$$E = \iint_{\Omega} \mu \|\epsilon\|_F^2 + \frac{\lambda}{2} \text{tr}^2(\epsilon) \, dx,$$

where $\epsilon = (\nabla v + \nabla v^\top) / 2$ is the infinitesimal strain tensor and $\mu, \lambda$ are Lamé's coefficients.

The Poisson’s Equation sees a lot of applications in the field of computer graphics. It can be used to reconstruct shapes from different domain [KBBH06], to find the potential component of a vector field [TLHD03], or to find smooth function extensions from boundary values [SHL*07]. The solution to the Poisson’s equation is the minimizer of the Dirichlet energy of a scalar function $f$:

$$E = \iint_{\Omega} \|\nabla f\|^2 \, dx.$$

For both the linear elasticity and the Poisson’s problem, their induced Euler-Lagrange equations are elliptic PDEs. Consequently, they can be discretized into a linear system $Ax = b$ where the left-hand side is the stiffness matrix. To efficiently extract the minimal and maximal non-trivial eigenpairs of a large sparse matrix $A$, we optimize the Rayleigh quotient $x^\top Ax / x^\top x$ using the nonlinear conjugate gradient method (NCG) [GSPF92] with approximate inverse preconditioner [BMT96]. This eigensovor converges very efficiently when the number of to-be-computed eigenpairs are small.

The Stokes Problem approximates creeping flow where advective inertial force can be ignored. This problem can also find applications in computational fabrication [ZYZZ15] and physics based animation [BB10]. Similarly, following FEM, the step of divergence free projection in splitting scheme minimizes a component-wise Dirichlet energy under the incompressible constraints:

$$E = \iint_{\Omega} \|\nabla v_x\|^2 + \|\nabla v_y\|^2 + \|\nabla v_z\|^2 \, dx$$

s.t. \( \nabla \cdot \mathbf{v} = 0 \) \quad \text{on} \quad \partial \Omega,$$

where $\mathbf{v}$ is the velocity field. We use Dirichlet boundary condition and fix boundary velocities to $\mathbf{v}$. This system can still be considered
L.E. Stability

Poisson Stability

Stokes Stability

L.E. Accuracy

Poisson Accuracy

Stokes Accuracy

Figure 8: Metrics are ranked based on their correlations with the Stability (left column) and Accuracy (right column) measurements of the linear elasticity problem, Poisson’s problem, and Stokes problem, respectively.

as an elliptic PDE in the subspace specified by the constraints. We use the same discretization setting as described earlier. Therefore, the resulting system is a saddle point linear system:

\[
\begin{pmatrix}
A_v & B^T \\
B & 0
\end{pmatrix}
\begin{pmatrix}
v \\
p
\end{pmatrix} = \begin{pmatrix}
b_v \\
0
\end{pmatrix},
\]

where \(A_v = A \otimes I^{1 \times 3}\) is the component-wise Laplace operator, \(B\) is the divergence operator, \(p\) is the Lagrangian multiplier and \(b_v\) is the boundary condition. The system can be solved only for the primal variables using the Schur complement matrix:

\[
A - A^{-1}B(BA^{-1}B^T)^{-1}BA^{-1}
\]

where \(A\) is the left-hand side. It is the non-trivial eigenpairs of this reduced system that we care about. To extract these eigenpairs while avoid forming this dense system directly, we instead use \(A_v\) as the left-hand side and project each search direction \(d\) of NCG onto the constrained subspace by solving:

\[
\arg\min_{d^*} \|d^* - d\|^2 \quad \text{s.t.} \quad Bd^* = 0,
\]

to get an equivalent result [KGW00]. Equation 3 can be solved very efficiently by pre-factorizing the dual system: \(BB^T\) using Cholmod [CDHR08]. \(d^*\) is then found from: \(d^* = (I - B^T(BA^{-1}B^T)^{-1}B)d\).

4.2.2. Accuracy and Stability of Elliptic PDEs

In the following, we show that the eigenvalues of the discretized system \(Ax = b\) are effective indicators of not only the hex-mesh quality, but also the discretization error of the elliptic PDEs solved on the continuous and the discretized domains, based on which we derive our eigenvalue-based accuracy and stability measurements for the elliptic PDEs.

Eigenvalues, \(\{\lambda_i\}\) are widely used as indicators of the quality of discretization of elliptic PDE on a certain mesh and the conditioning and stability of the discretized linear system. As explained in [BPM*95], any discretization scheme introduces additional numerical stiffness because of the extra forces needed to restrict a continuous solution to a small subspace dictated by the mesh. The worse the discretization is, the stronger these restrictions are, and the “stiffer” the linear system is. As a result, the minimal non-zero eigenvalue, \(\lambda_{\min}\), indicates the accuracy of a discretization on a certain mesh, as compared to the continuous case [She02]. Note that this smallest non-zero eigenvalue indicates the \(\ell_2\) error of the approximation, which is the most common error measurement used.
by many PDE problems [She02]. The accuracy of the problems performing on a given discretized hex-mesh is computed as the difference of the smallest non-zero eigenvalue with the ground truth. As the ground truth is typically not available, we approximate it using the minimum eigenvalue obtained from the same simulation carried out on a dense (e.g., 100K elements for all objects) hex-mesh of the same object.

The other metrics we will use are the condition number and the maximal eigenvalue $\lambda_{\max}$. Unlike $\lambda_{\min}$ which measures the accuracy of the discretization itself, these two metrics measure the computational efficiency and stability of an already discretized linear system. The condition number of a symmetric linear system (under $\ell_2$ norm) is determined by $\lambda_{\max}/\lambda_{\min}$. Since the convergence speed of a typical iterative linear system solver relies on this number [Saa03], the condition number measures the computation efficiency of the discretized linear system. Meanwhile, the maximal eigenvalue $\lambda_{\max}$ is a robust indicator of hex-elements with bad quality.

### 4.2.3. Evaluation using Eigenvalue-based Measurements

We can now investigate the impact of the hex-mesh quality metrics on the above three applications by including the Accuracy and Stability metrics to the correlation matrix. Similar to the above application independent study, we compute the average matrix of the correlation matrices for the individual datasets, as shown in Figure 7. Again, the deviation matrix (in supplementary material) shows small variation across the correlation matrices of different objects. Based on the absolute values of their correlation coefficients with Accuracy and Stability, we rank all the existing metrics as shown in Figure 8. We color the histograms based on the clusters (see Figure 5) that the individual metrics belong to.

From these results, we have the following observations:

1. All average metrics have stronger correlations with both the Accuracy and Stability than those minimum/maximum metrics. While this phenomenon seems to contradict with the statement in [She02] that a single bad element can result in a very bad conditioning, it is not well-defined how “bad” the worst element of a hex-mesh should be for it to be considered as a bad element. It is possible that the quality of elements of the meshes used in our study is not sufficiently “bad” to significantly influence the conditioning, and the percentage of the “bad” meshes in the whole database is not high enough. We wish to explore this in the near future.

2. Among those metrics, MAF. Avg. has the strongest correlation with the Stability for all three applications and DM. Avg. ranked second. MAF (mean aspect Frobenius) is the average of the aspect Frobenius measured at the eight corners and the center of a hexahedron. For Accuracy, while MAF. Avg. has the strongest correlation with the Poisson’s and Stokes simulations, DM. Avg. tops the others for the linear elasticity simulations (with MAF. Avg. ranked 2nd). DM (i.e. Dimension) was specifically designed in the context of Sandia’s Pronto code, for stable time step calculation [SEK07]. It is calculated as the ratio of the element volume and its discrete gradient (see the supplemental document). Nonetheless, the ranking difference among all average metrics is almost negligible, coinciding with the discovery made in the application independent study (i.e. all average metrics have similar capability in measuring the overall quality of a hex-mesh).

3. The correlations of the metrics (especially average metrics) with the Stability are much stronger than their correlations with the Accuracy in the Poisson’s and Stokes equation solving applications (Figure 8(b,c)), while this is not the case for the linear elasticity problem, as all average metrics have strong correlation with the Accuracy metric (Figure 8(a)). This means that a hex-mesh with not-so-good overall quality will usually lead to slow numerical performance, as rigorously shown in [Saa03] for various iterative solvers, but has little impact to the accuracy of the Poisson’s and Stokes equation solving.

4. SKEW Max. that measures the orthogonality of the principal axes of a hexahedron, has the highest ranking among all Min/Max metrics for most applications, while the well-known S. Jac. Min is ranked much lower in all experiments. In fact, all other metrics from cluster 3 (Section 4.1) (i.e. in the same cluster as SKEW Max.) are all ranked higher than S. Jac. Min. By analyzing the formulation of these metrics (see supplemental document), we find that they all directly or indirectly characterize the conditioning of the element. For instance, SKEW takes the maximum absolute value of the dot products of pairs of the principal axes of a hexahedron. The closer the hexahedron to a regular cube, the smaller the SKEW is (i.e. all dot products tend to zero), indicating better conditioning of $J^TJ$ (i.e. all its eigenvalues tend to 1). If SKEW is large, at least one pair of the principal axes is far away from the orthogonal configuration, leading to large condition number of $J^TJ$ (i.e. $\lambda_{\max}/\lambda_{\min}$ is much larger than 1). Similarly, MER (maximum edge ratio) that computes the maximum ratio between the lengths of pairs of principal axes, TAP that measures the maximum ratio of a cross-derivative to its shortest associated principal axis, and DIAG that is the ratio of the minimum diagonal length to the maximum diagonal length, also characterize the conditioning of $J^TJ$ in various ways (see the supplemental document). In contrast, S. Jac. – the determinant of the normalized $J$ does not measure the conditioning of $J^TJ$, thus, it is ranked much lower than the above four metrics. On the other hand, the volume metric (i.e. VOL. Max.) that is computed as the product of the magnitude of the three principal axes of the hexahedron is ranked much higher than S. Jac., indicating that the scale information of the element does matter to the simulation quality, while S. Jac. is scale independent.

### 4.2.4. Evaluation using An Analytical Application

For the Poisson’s problem, one can define an arbitrary analytic solution $u_{\text{ref}}$, and reconstruct it on a certain hex-mesh with a given orientation. This is achieved by minimizing:

$$E = \int_{\Omega} \left\| \nabla f - \nabla u_{\text{ref}} \right\|^2 d\Omega,$$

and finally measuring the error using $L_2$ norm as $\varepsilon = \sqrt{\int_{\Omega} (u_{\text{ref}}(x,y,z) - f(x,y,z))^2 d\Omega}$. Note that other error measurement can be used here than the $L_2$ error. In our experiment, we use $u_{\text{ref}}(x,y,z) = \cos(x) + y \times z + \exp(x + z)$, where the domain of $(x,y,z)$ is varying for different datasets. The correlations (shown in the first row of the correlation matrix in Figure 7) between the
quality metrics with the accuracy of the solution is mostly consistent with the eigenvalue based evaluations (i.e., the top two rows exhibit similar patterns). The metric ranking is shown in Figure 9, in which DM, Avg., S JAC, Avg., SP, Avg. and MAF Avg. exhibit strong correlations with the accuracy of the computation, which largely coincides with our observation for the Accuracy metric of the Poisson’s equation solving (Figure 8, second row, right). Such analysis can also be done for Stokes problem using established analytical solutions, e.g., a flow computed within a sphere [Bat07].

Figure 9: Metrics ranking for the analytical application of Poisson’s equation solving.

Conclusions of our application dependent study: Overall, the above results and analysis could lead to a potentially useful guideline, at least for our generated datasets. That is, in order to achieve a more accurate and stable computation in solving elliptic PDEs, the quality of all elements, measured by those average metrics, should be the center of the meshing process once the mesh is inversion-free. Also, metrics that characterize the conditioning of the elements, e.g., MAF Avg., DM Avg., SKEW Max., MER Max., and TAP Max., usually have stronger correlations with the quality of solving elliptic PDEs than other metrics, and thus, should be the quality that the meshing techniques try to optimize. In addition, based on our observation in the application independent study, those metrics have strong correlations with all average metrics (Figure 6(c)). Therefore, one will expect that improving any one of these metrics would lead to the improvement of the overall mesh quality measured by other metrics.

5. Conclusion

This work provides a statistics based solution to a challenging and important metric evaluation problem that has never been addressed before, i.e., the evaluation of the effectiveness of the quality metrics for a hex-mesh as a whole regarding different applications. This analysis framework can be applied to the quality metric evaluation of other types of meshes (e.g., tet-meshes), which we believe is an important contribution to the meshing community. With our analysis framework, important and useful findings are discovered based on our carefully generated mesh datasets, including the clustering structures of metrics and the importance of the average quality metrics to elliptic-based applications that are not well understood before. We believe our analysis framework and the obtained findings represent a sturdy step toward the comprehensive understanding of the effectiveness of the individual quality metrics, which is much needed by the meshing community.

There are a number of limitations of this work.

Authenticity of the datasets: Similar to previous evaluation works [MHKSZ01, MN11] that perform the analysis using artificial datasets, all our discoveries are derived from the analysis on datasets generated by the proposed perturbation strategy. There is no guarantee that these findings accurately imply the behaviors of the quality metrics when considering real hex-meshes used in practice. However, this should not demolish the value of the proposed statistics based analysis framework, as it does not depend on the datasets used. On the other hand, obtaining a more realistic hex-mesh dataset for the proposed evaluation is notoriously challenging given the criteria listed in Section 3, which should be addressed in the future to further verify our current discoveries.

Sampling: The current sampling strategy does not explicitly control the maximum metrics, some artificial clusters can be observed. See the inset figure for an example. Since the analytical relations among the quality metrics are not well-studied, it is hard to evaluate what is a “good” sampling, especially considering that we attempt to generate samples in a high dimension space (each metric represents a dimension).

Connectivity: The extreme metrics measure the highest distortion of a specific element among all elements in the hex-mesh. As shown in the inset, this element may be at different locations (i.e., both geometrically and topologically) of the mesh, as shown in the inset (i.e., the colored hexahedra), which may partially explain the fluctuations of the correlations among various metrics and simulations (see the deviation matrices shown in Figures 1 and 2 in the supplemental material). A more detailed study on how different local configurations affect the correlation study may be needed in the future.

Discretization:
The inset figure shows six hex-meshes of a cube with increasing resolutions (from the left to the right). Similar to the above studies, we generate one dataset for each resolution using the proposed perturbation strategy. Their respective application dependent correlation matrices are shown in the inset, which are transposed compared to Figure 7. From the visual comparison, we observe that the correlations among metrics and simulations vary for different numbers of elements. In addition, the patterns of the individual rows in the shown correlation matrices are not consistent, which requires a detailed analysis to understand its cause.

Mesh types: While the proposed framework is designed to evaluate quality metrics from purely conformal all-hex meshes, quality
metrics for other types of meshes, for example, hybrid meshes and T-meshes, can be evaluated similarly.

Acknowledgments

The authors wish to thank the anonymous reviewers for the constructive comments. This work was supported in part by NSF (IIS 1553329 and IIS-1524782), and NSFDC (61522209 and 61210007).

References


[Ins85] Inselberg A.: The plane with parallel coordinates. The Visual Computer 1, 2 (1985), 69–91. 4


[Joe08] Joe B.: Shape measures for quadrilaterals, pyramids, wedges, and hexahedra, 2008. 1, 2


