Subdivision Isogeometric Analysis: A Todo List

Abstract

Subdivision techniques are natural when extending isogeometric analysis beyond regular grid control nets, i.e., to integrated design and analysis of irregular layouts in several variables. This report identifies and discusses several challenges that need to be addressed to place the use of subdivision techniques for isogeometric analysis on a solid foundation. The report highlights issues of regularity, approximation and numerical quadrature. The challenges are illustrated by an implementation solving solving elliptic partial differential equations in the space of trivariate Catmull-Clark subdivision.

1 Introduction

Due to their built-in ability to refinably model smooth functions with a non-regular control net, subdivision surfaces have repeatedly been considered as basis functions both for modeling geometry and solving differential equations on such a geometric domain, see e.g. [COS00, BHU10, NKP14, XDD’20]. Many publications, see Section 7, have explored the use of bivariate subdivision algorithms for solving elliptic PDEs like Poisson’s equation or the biharmonic equation on surfaces and reported numerical convergence rates in the range from $O(h^2)$ to $O(h^4)$. However, a closer look at the theoretical underpinnings reveals significant gaps in our understanding and in our arsenal of algorithms. Filling these gaps is an indispensable prerequisite for the further development of subdivision as a unified tool for modeling and analysis.

The purpose of this report is therefore to identify missing pieces in the foundation of subdivision-based simulation and to comment on directions to be explored—but not to present solutions. We hope that this will stimulate discussions and advance the future development of the field. Some of the thoughts presented here in some detail can be found already in [RS19], but they appear there only as a rough sketch as part of a much broader scope.

Many of the subsequent statements and remarks have the status of hypotheses rather than proven facts and reflect our personal opinions based on longstanding experience with subdivision methods. The arguments are addressed to old and new experts in the field. That is, we assume that the reader is familiar with basic concepts of the subdivision framework so that we can succinctly approach the various topics. If this is not the case, the text book [PR08] may serve as an introduction to the field. Also the notation used here is adopted from this reference.

We focus on isogeometric analysis (IGA) for analyzing an elliptic partial differential equation (PDE) of the form $Lu = f$ on some domain $\Omega \subset \mathbb{R}^d$. IGA, a term proposed by [HCB05], uses basis functions $b_k$, $k \in K$, which are not defined directly on $\Omega$, but on a set $S$, typically a collection of simple geometric domains such as cubes, squares, or triangles. A parametrization $\varphi : S \rightarrow \Omega$ is constructed as a linear combination of the basis $b_k$. The approximation $u_h \approx u$ of the solution is then sought in the form $u_h = z \circ \varphi^{-1}$ and one has to solve a system of equations to determine the coefficients $z_k$ of $z = \sum b_k z_k$. Defining basis functions $\beta_k := b_k \circ \varphi^{-1} : \Omega \rightarrow \mathbb{R}$ and the Galerkin projection of the partial differential equation, $u_h = \sum \beta_k z_k$ needs to be sufficiently regular: For elliptic problems of order $2m$, one needs functions in the Sobolev space $H^m(\Omega)$, which is the space of all functions on $\Omega$ whose weak partial derivatives of order up to $m$ are square integrable.

For second order problems, i.e. $m = 1$, the requirement $\beta_k \in H^2(\Omega)$ is easy to fulfill. For instance, one can choose functions $b_k$ that are piecewise polynomial and continuous. In particular, piecewise linear approximations by so-called hat functions over a triangulation of $\Omega$ are commonly used in Finite Element Analysis (FEA). Also popular are piecewise polynomials of degree $p > 1$. For fourth order problems, i.e. $m = 2$, the requirement that $\beta_k \in H^4(\Omega)$ is less easy to achieve. Consider the parameterized surface $s(y) = (\varphi(y), z(y)), y \in S$. The trace $s(y) : y \in S$ equals the graph $\{(x, u_h(x)) : x \in \Omega\}$ of $u_h$. However surface representations that are barely continuous, like those based on hat functions, produce traces that are not smooth enough to be graphs of $H^2$ functions. Hence, ignoring for now the fine print of Sobolev embedding theorems, it is sensible to assume that the basis $b_k$ generates $C^1$ surfaces since this ensures $\beta_k \in C^1$. $C^1$ smoothness is neither necessary nor sufficient, but a good starting point when searching for suitable bases. As a consequence of this insight, there is a growing interest of the IGA community in modifying or directly applying smooth surface constructions developed decades ago for applications in Computer Aided Geometric Design.

Of particular interest are three surface representations for meshes with irregularities. In the classification of [Pet19]:

- Geometrically continuous splines;
- Singular corner parametrization splines;
- Singular corner parametrization splines;

\[ a^{\text{Technische Universität Darmstadt, Dept. of Mathematics, Darmstadt, Germany.}} \]
\[ b^{\text{University of Florida, Dept. CISE, Gainesville, USA.}} \]
\[ c^{\text{Numerical Geometry Ltd., Cambridge, UK.}} \]
Subdivision splines.

Geometric continuity [DeR90, Pet02] is a well-developed concept for the representation of $C^1$ and even $C^2$ surfaces and can certainly be used for bivariate problems. However, a generalization to the trivariate case, which poses the real challenge, seems to be extremely complicated. [BJM18, BK19] join just one pair of tri-linearly parameterized face-adjacent boxes and [KV22] consider tri-linearly parameterized multi-patch volumes with exactly one inner (irregular) edge. The challenge is the complicated interaction of reparametrizations surrounding an irregular point. A systematic disadvantage of function spaces based on geometric continuity is a missing natural refinement property, as provided by the next two spline classes.

Singular parametrization [Pet91, Rei97, MR22] is slightly less attractive for geometric modeling since the visual quality typically suffers from unbounded curvature. However, it may well be useful for simulation, and first steps have been taken in that direction [TSH17]. Generalization to the trivariate case is possible [Pet20], but regularity assumptions need to be scrutinized in detail [TJ12, Rei22].

Subdivision is the method of choice in Computer Graphics since it combines conceptual simplicity with smooth curved shape. A few trivariate algorithms are known, commending a closer look since several theoretical and practical aspects are scattered. Below, we list a series of open problems that we consider central.

In the next section, we comment on subdivision algorithms for surface generation and possible generalizations to the trivariate case. Here, we recommend to develop algorithms with an optimized spectrum and also to develop a trivariate variant of Doo-Sabin subdivision. In Section 3, we discuss regularity issues and make clear that we know more or less nothing about trivariate schemes in that respect. Filling that gap is perhaps the most essential and the most challenging task ahead. Approximation issues are addressed in Section 4. Technical, but nevertheless important aspects of numerical quadrature for the compilation of Ritz-Galerkin systems are considered in Section 5. Section 6 addresses the conditioning of the linear systems. Section 7 presents results of an implementation for the simulation of trivariate partial differential equations. Convergence lower than for regular splines points to the need for improving our understanding of the theoretical background.

2 Algorithms

Many subdivision schemes on regular grids are available. However, in the context of IGA, they offer no advantages over explicit representations such as tensor product B-splines. The power of subdivision becomes apparent where non-regular meshes are needed. For bivariate surfaces in three-space, the most popular linear subdivision algorithms are, for quad meshes, Catmull-Clark subdivisions [CC78] and, for triangular meshes, Loop subdivision [Loob87]. These schemes generate $C^1$ surfaces that are $C^2$ everywhere except at extraordinary points (EOPs). Moreover, the generated surfaces are infinitely piecewise polynomial. Various modifications of the original schemes have been proposed in the past, see for instance [ADS06, MM18], mostly with the goal to improve the visual appearance by tuning towards 1 the ratio $\mu/\lambda^2$ of the subdominant eigenvalue $\mu$ of the subdivision process to the squared subdominant eigenvalue $\lambda^2$. Another aspect, called polar artifact [BK04], is equally important. It denotes the deviation of $\lambda$ from the regular shrinkage factor $\lambda_0$. For Catmull-Clark, Loop, and all other binary schemes, $\lambda_0 = 1/2$.

Figure 1 illustrates the polar artifact for Catmull-Clark. For valence $n = 3$, we have $\lambda < 1/2$, corresponding to a higher rate of contraction than in the regular parts. By contrast, valence $n = 9$ yields $\lambda > 1/2$, corresponding to slower shrinkage.

Task 2.1 should be understood as a call to study the apparently not well-known literature and to promote and advance the results and methods presented therein.

Task 2.1. Study and employ variants on Catmull-Clark and Loop with small polar artifact $|\lambda - \lambda_0|$ and smallest possible $\mu$.

The task could be extended to other schemes, but Catmull-Clark and Loop deserve special attention because they are so widespread. At issue is that values $\lambda > \lambda_0$, as encountered for instance for standard Catmull-Clark subdivision and valency $n \geq 5$, lead to a locally slower decay of the mesh size. A desired level of local resolution may therefore require more levels of refinement. Concerning $\mu$, it is known that the principal curvatures are bounded if and only if $\mu \leq \lambda^2$. Even larger values yield square integrable principal curvatures [RS01], as required for the applications we have in mind. But we will point out in Section 5 that the numerical integration of unbounded functions may cause severe problems. In Computer Graphics, the main

---

1Of course, many more schemes are available, for instance Doo-Sabin, Butterfly, Velho 4-8, $\sqrt{3}$, or Simplest subdivision, see e.g. [PR08].

2The most popular non-binary scheme is $\sqrt{3}$-subdivision, whose regular shrinkage factor $\lambda_0 = 1/\sqrt{3}$ explains its name.
area of application of subdivision today, values $\mu < \lambda^2$ are not desirable because they inevitably lead to unaesthetic flat spots. However, in the context of approximation, such considerations are not relevant. In Section 5 and Section 4, we will discuss the benefits of our recommendation concerning $\mu$: smaller is better.

Subdivision surfaces can be used to approximate solutions of fourth order partial differential equations in two variables. However, it must be doubted if this approach has really the potential to become the method of choice. Other $C^1$ elements, like Argyris, Clough-Tocher, or Powell-Sabin, consisting of finitely many polynomial pieces, are well established and not easily outperformed by subdivision techniques.

The situation is different in three (or more) variables. Here, $C^1$ elements are much harder to construct. Splits of macro tetrahedra into smaller ones serving that purpose are known [AS05a, AS05b, Sor09, Spe13], but are typically quite complicated. So trivariate subdivision may offer a competitive solution. Considering hex meshes only (there are no regular tetrahedral meshes) and disregarding their precise technical specifications for partitions without gaps and overlaps, and their challenging construction (see for instance [ZVC20] and the references therein), we distinguish

- regular edges, shared by four cubes;
- regular vertices, shared by six regular edges;
- extraordinary edges, shared by more or less that four cubes;
- semi-regular vertices, shared by two extraordinary and several regular edges;
- extraordinary vertices, shared by more than two extraordinary edges.

A natural way to design a trivariate subdivision scheme is to use

- standard subdivision rules for trivariate cubic B-splines on regular parts of the mesh;
- the tensor product of univariate cubic B-spline subdivision and Catmull-Clark near extraordinary edges;
- special rules near extraordinary vertices.

Actually, the algorithm suggested in [BSWX01] proceeds that way, while the trivariate scheme [MJ96] does not simplify the construction around extraordinary edges into the extruded bivariate case. Interpolatory trivariate schemes are presented in [CMQ03] and [XXD+20]. One problem with all these algorithms is the structure of the spectrum of the subdivision matrix representing subdivision near an extraordinary vertex. Let the eigenvalues be ordered according to

$$\lambda_0 = 1 > |\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \geq \cdots .$$

Then the eigenvectors $v_1$, $v_2$, $v_3$ corresponding to the three subdominant eigenvalues $\lambda_1$, $\lambda_2$, $\lambda_3$ define a trivariate characteristic map, analogous to the well-known construction in the bivariate case. Asymptotically, near an extraordinary vertex, iterated refinement of a subdivision volume is akin to repeated scaling of this map by the factor $\lambda_i$ in the $i$th coordinate direction. However, the subdominant eigenvalues are all different unless the configuration is highly symmetric. This results in a loss of isotropy in the sense that different directions are scaled down at different rates. Under repeated refinement, the $v_2$-component decays faster than the $v_3$-component, which in turn decays faster than the $v_1$-component. In descriptive terms, this means that the configuration becomes flatter and flatter and then also slimmer and slimmer so that it is eventually squeezed to an elongated, line-like object. This unbalanced behavior is certainly unwanted for reasons very similar to those discussed for Task 2.1. Figure 2 illustrates the phenomenon. Given an extraordinary point where two regular and four extraordinary edges meet and using the algorithm of [BSWX01], the control mesh of the characteristic map is shown on the left hand side. The middle and right column show the mesh after 7 and 13 iterations, respectively. The leading eigenvalues are

$$\lambda_0 = 1, \quad \lambda_1 \approx 0.5604, \quad \lambda_2 \approx 0.4692, \quad \lambda_3 \approx 0.4326, \quad \lambda_4 \approx 0.3177.$$  

This leads to the next task.

**Task 2.2.** Construct a trivariate generalization of Catmull-Clark with triple subdominant eigenvalue $\lambda = 1/2$ and smallest possible $\mu$.

The challenge posed by Task 2.2 is much harder than the corresponding task in the bivariate case: While all bivariate configurations are closely related in the sense that they differ only by the valency of a single extraordinary vertex, the number of combinatorially dissimilar situations is unbounded in the trivariate case. This makes it difficult to develop a unified analytical framework beyond the study of a list of special cases.

An interesting alternative, seemingly not considered so far, is the construction of a trivariate dual scheme generalizing Doo-Sabin [DS78]. While the lack of curvature continuity between neighboring patches limits applications of Doo-Sabin surfaces in geometric modeling, it may still be a competitive approach to solving approximation problems. We will discuss aspects of approximation in Section 4 in more detail, but it seems to be reasonable to expect higher order convergence for Doo-Sabin type algorithms than for standard piecewise linear elements. Furthermore, Doo-Sabin yields smaller bandwidths of Galerkin matrices than Catmull-Clark. This leads to faster matrix compilation and also suggests a faster solution of the resulting linear systems. Thus, it may be worthwhile to consider

**Task 2.3.** Construct a trivariate generalization of Doo-Sabin with triple subdominant eigenvalue $\lambda = 1/2$ and smallest possible $\mu$.

The tensor product of standard Doo-Sabin and univariate quadratic B-spline subdivision can be used to treat $n$-sided prisms, which are the analogue of extraordinary edges, as discussed above. Instead of extraordinary vertices, one has now to deal with general polyhedra that the algorithm has to map to smaller polyhedra of the same combinatorial structure. First experiments...
indicate that a triple \( v_1, v_2, v_3 \) of suitable subdominant eigenvectors can be found among the eigenvectors of the adjacency matrix of the graph corresponding to the vertices and edges of the polyhedron. Together with the vector \( e = [1, \ldots, 1] \), the matrix

\[
A = \frac{ee^t}{\|e\|^2} + \frac{1}{2} \left( \frac{v_1 v_1^t}{\|v_1\|^2} + \frac{v_2 v_2^t}{\|v_2\|^2} + \frac{v_3 v_3^t}{\|v_3\|^2} \right)
\]

seems to be a good candidate for the mapping of the polyhedron. At least, this matrix \( A \) has the proper spectrum with a triple subdominant eigenvalue \( \lambda = 1/2 \).

3 Regularity

The analysis of linear subdivision schemes in several variables over a regular lattice has reached a state of maturity, see for instance [CDM91, DL02, Sab10]. Also bivariate subdivision surfaces in \( \mathbb{R}^3 \), with extraordinary points, are well understood [Rei95, PR08]. In the following, we use methods and notation from [PR08].

A linear \( d \)-variate subdivision scheme produces limit functions \( x : \mathcal{S} \to \mathbb{R}^s \) from control points \( x_k \in \mathbb{R}^s, k \in K \). The domain has the form \( \mathcal{S} = \Sigma \times I \), where \( I \) is some finite index set and the base cell \( \Sigma \subset \mathbb{R}^d \) is chosen as follows: In the bivariate case \( d = 2 \), the base cell is either a square or a regular triangle. In the trivariate case \( d = 3 \), the base cell is either a cube or a regular tetrahedron. Higher dimensional cases are conceivable, but not considered here. Denoting the generating functions by \( b_k \), the function \( x \) can be written in the form \( x = \sum_{k \in K} b_k x_k \). In the following, scalar functions \( (s = 1) \), parameterizations \( (s = d) \), and hyper-surfaces \( (s = d + 1) \) are of special interest.

In the case of bivariate surfaces in \( \mathbb{R}^3 \) \((s = d + 1 = 3)\), necessary and sufficient conditions for \( L^p \) integrability of the principal curvatures of the limit surfaces are known [RS01]. In that paper, it is shown that the necessary and sufficient \( C^1 \) conditions for linear, stationary subdivision schemes are strong enough to guarantee square integrability of the principal curvatures with respect to the surface measure.

In the applications we have in mind, such subdivision schemes are exploited for the approximation of elliptic fourth order PDEs like \( \Delta^2 u = f \), defined on some domain \( \Omega \subset \mathbb{R}^2 \), in the sense of isogeometric analysis. That is, one uses subdivision functions \( \varphi := \sum b_k \varphi_k : S \to \Omega \) and \( z := \sum b_k z_k : S \to \mathbb{R} \) to approximate the solution \( u \) by \( u_h := z \circ \varphi^{-1} \), where the control points \( \varphi_k \subset \mathbb{R}^2 \) are determined in advance such that the parameterization \( \varphi : S \to \Omega \) is bijective, and the coefficients \( z_k \in \mathbb{R} \) are characterized by the weak formulation of the PDE. If the surface \( x = [\varphi, z] \) is normal continuous, then \( u_h \in C^1 \), which can be concluded from

\[ \int_{\Omega} |\nabla^2 u_h |^2 \, d\mathbf{s} \leq C \int_{\Omega} |\nabla^2 f |^2 \, d\mathbf{s} \]
continuity of $\varphi^{-1}$ and
\[
n = \frac{\partial_1 x \times \partial_2 x}{\|\partial_1 x \times \partial_2 x\|} = \frac{1}{\sqrt{1 + \|\nabla u_\varphi \circ \varphi\|^2}} \begin{bmatrix} \nabla u_\varphi \circ \varphi \\ -1 \end{bmatrix}.
\]
Furthermore, square integrable principal curvatures of $x$ suggest $u_\varphi \in H^2$, i.e., square integrability of the second partials of $u_\varphi$. However, we are not aware of a formal proof of this conclusion. The fact we are interested in is not based on the special structure of subdivision surfaces, but applies to general surfaces that can be expressed as a graph of a function.

**Task 3.1.** Prove the following conjecture: Given a normal continuous surface $x = [\varphi, z]$ with $\varphi$ invertible and principal curvatures $\kappa$, square integrable with respect to the surface measure, i.e.,
\[
\sum \int_X \kappa^2 \, dx < \infty,
\]
then $u_\varphi = z \circ \varphi^{-1} \in H^2$.

The statement may seem plausible, but it is certainly not trivial. For instance, one can show by a counterexample that continuity of the principal curvatures does not imply continuity of the second partials of $u_\varphi$, since the principal directions may be discontinuous. It is also well possible (but not likely in our experience) that a corresponding theorem can be found in the literature in the field of classical differential geometry.

In contrast to the bivariate case, the smoothness analysis of trivariate schemes is completely unsettled. Following the standard setup of the bivariate case, a trivariate subdivision surface $x : S \to \mathbb{R}^4$ can be defined on a domain of the form $S = \Sigma \times I$, where $\Sigma := [0, 1]^2$ is the unit cube and $I$ is a finite index set. Note that the image space is now four-dimensional in accordance with the fact that the graph of the trivariate function $u_\varphi = z \circ \varphi^{-1}$ is a hyper-surface in $\mathbb{R}^4$. Now, at least in the case of well-defined and linearly independent partial derivatives, the normal vector is given by
\[
n := \frac{\hat{D}^x}{\|\hat{D}^x\|},
\]
where $\hat{D}^x$ is a vector perpendicular to all three partials of $x = [x_1, x_2, x_3, x_4]$. Denoting the standard basis in $\mathbb{R}^4$ by $e_1, \ldots, e_4$, such a vector can be computed by the formal determinant
\[
\hat{D}^x := \det \begin{bmatrix} e_1 & e_2 & e_3 & e_4 \\ \partial_1 x_1 & \partial_2 x_1 & \partial_3 x_1 & \partial_4 x_1 \\ \partial_2 x_1 & \partial_2 x_2 & \partial_3 x_2 & \partial_4 x_2 \\ \partial_3 x_1 & \partial_3 x_2 & \partial_3 x_3 & \partial_3 x_4 \end{bmatrix}
\]
Normal continuity of the subdivision surface $x$ means that there exists a well-defined limit of the normal vector when approaching any point on an extraordinary edge or an extraordinary vertex. The corresponding task is the following:

**Task 3.2.** Formulate and prove a theorem providing sufficient conditions for normal continuity of a trivariate subdivision algorithm.

Building the algorithm from a smooth surface scheme like Catmull-Clark subdivision and a smooth univariate scheme along an extraordinary edge does imply normal continuity there. The proof is not difficult, but, according to our experience, also slightly more involved than the seeming obviousness of the result suggests. Even more challenging is the situation at an extraordinary point. It is quite clear that the structure of the spectrum of the subdivision matrix $A$ will be crucial as well as the properties of a trivariate characteristic map, defined by the subdominant eigenvectors $v_1, v_2, v_3$. But what exactly is the formulation equivalent to the bivariate case requiring regularity of the characteristic map? Now, ordinary and extraordinary edges meet, and each extraordinary edge is special itself. Or, to put it differently, the characteristic map is now composed of infinitely many regular pieces. The implications of this fact are not easy to anticipate.

The second step of the analysis concerns single-sheetedness of trivariate subdivision surfaces, preventing local self-intersections. As in the bivariate case, this property is necessary to construct an injective parametrization $\varphi : S \to \Omega \subset \mathbb{R}^3$. This time, a theorem is at least easy to formulate.

**Task 3.3.** Show that a normal continuous subdivision scheme generates locally single-sheeted surfaces for almost all input data if the characteristic map is injective.

The restriction of the statement to almost all input data means the following: As in the bivariate case, single-sheetedness cannot be expected for all choices of initial control points. However, the set of exceptional configurations should be of measure zero.

A subdivision algorithm is called $C^1$ if it typically generates surfaces that are normal continuous and single-sheeted. Once sufficient conditions in this regard are established, the next step is

**Task 3.4.** Show that trivariate subdivision algorithms, such as 3d variants on Doo-Sabin or Catmull-Clark, are $C^1$.

A crucial step on the way will be the eigenanalysis of the subdivision operator. In this respect it should be noted that it is unclear if one can generalize, to the trivariate setting, the discrete Fourier transform (see [PR08], Sect. 5.4), which is so helpful when dealing with symmetric bivariate schemes. Discrete spherical harmonics, which can be thought of as a higher-dimensional analogue of the discrete Fourier transform, may be worth an investigation.

Notably, it can no longer be expected that all possible combinatorial configurations can be treated by a single analytic approach. For bivariate schemes this is possible since the valency of an extraordinary point simply appears as a parameter in an otherwise uniform setup. It should also be noted that symmetry plays a crucial role for the simplified verification of injectivity of the characteristic map of bivariate schemes. We do not expect anything similar for trivariate schemes.

After the $C^1$-analysis is completed, second order properties have to be investigated:
Task 3.5. Find conditions guaranteeing square integrability of the principal curvatures of trivariate subdivision surfaces and verify them for selected algorithms.

Compared with the previous tasks, this one seems less complicated. Concretely, we conjecture that schemes satisfying the $C^1$-conditions will always produce surfaces with square integrable principal curvatures, as in the bivariate case. Further, we expect that the last step concerning $H^2$-regularity of the graph parametrization $u_h = z \circ \varphi^{-1}$ of a trivariate subdivision surface $x = (\varphi, z)$ will need no further effort beyond Task 3.1 since the techniques developed for that purpose should be applicable in arbitrary dimensions.

4 Approximation

To put is mildly, our understanding of approximation properties of bi- and trivariate subdivision functions is incomplete. Even if all issues raised above are settled, the prospects of subdivision in IGA stand and fail with the ability to approximate the sought solution. According to Cea’s Lemma and standard duality arguments for elliptic, self-adjoint PDEs, the approximation generated by the Ritz-Galerkin method is, up to a constant, as good as the best possible approximation. That is, we need to turn our attention to

Task 4.1. Prove error estimates of the following type: Let $\varphi : S \to \Omega$ be a parametrization of the domain $\Omega$ and let $h$ be the maximal diameter of images of grid cells under $\varphi$. For each sufficiently smooth function $f : \Omega \to \mathbb{R}$ there exists a subdivision function $z : S \to \mathbb{R}$ such that

$$\max_{|i|=m} \|\partial^i(f - z \circ \varphi^{-1})\|_2 \leq c_{k,m} h^{m-k} \max_{|i|=m} \|\partial^i f\|_2,$$

where the orders $k, m \in \mathbb{N}_0$ and $c_{k,m}$ is a constant.

In [PXXX16] and [PRY21], the authors establish such an error estimate for the Catmull-Clark scheme and $m = 2$. However, it is not clear if this bound, which is only as good as that for piecewise linear elements, is optimal. Ideally, one would hope for $m = 4$, i.e., quartic rather than quadratic convergence, because cubic tensor product B-splines converge that fast and their subdivision generalization should preserve it. On the other hand, though far from being rigorous, the following argument casts doubt on this expectation.

For simplicity, let us consider the bivariate case with a single extraordinary point of order $n$ and standard Catmull-Clark subdivision. For all regular patches, we assume that the error $e = f - z \circ \varphi^{-1}$ is of the optimal order $O(h^4)$. However, for the $n$ patches sharing the extraordinary point, we know only that $z \circ \varphi^{-1}$ can represent linear functions, but not polynomials of higher order. Hence, locally, we expect the error to be only of order $O(h^3)$. We see no reason to believe that the value of the subsubdominant eigenvalue $\mu$ has any influence on the approximation order. In particular, the case $\mu = \lambda^2$, which desirable for well-shaped surfaces, does not by itself seem to offer any extra benefit for approximation in general.

The area of the image of a grid cell is $a = O(h^2)$, the number $N$ of regular patches is $N = O(h^{-2})$, and the number of irregular patches is $n$. The squared $L^2$ error $\|e\|_2^2$ can be estimated by multiplying the squared local maximal errors by the area $a$ and summing over all patches. We find

$$\|e\|_2^2 = (N O(h^4)^2 a + n O(h^3)^2 a)^{1/2} = (O(h^8) + O(h^6))^{1/2} = O(h^5).$$

This cursory inspection suggests that the reduced approximation power near the extraordinary point dominates the total error. Similar considerations apply in the trivariate case, where extraordinary edges contribute more to the error than extraordinary points. A better than predicted $O(h^3)$ convergence of interpolatory Catmull-Clark subdivision for Poisson’s equation with zero boundary values on a trivariate ball is reported in [XXD+20, Fig 7]. However, this measurement applies only to the first refinement step and the convergence graph flattens out to approximately $O(h^3)$ by the third refinement.

Assuming that the argument leading to (1) is conclusive, what can be done? First, adaptive refinement, which is mandatory for any serious implementation anyway, can be used to locally reduce the mesh size until the desired accuracy is achieved. The interesting question here is how many extra levels are necessary on average to compensate for the locally reduced approximation order. Second, variants of Catmull-Clark subdivision could be developed that locally reproduce quadratic functions. Then the local error $O(h^3)$ would lead to the global error $O(h^4)$, as wanted. However, this is more or less equivalent to the notorious problem of constructing $C^2$ subdivision schemes. Third, one could simply accept the result and be happy that $O(h^3)$ is still better than the typical $O(h^2)$ convergence provided by piecewise linear elements. However, in this case, replacing Catmull-Clark subdivision by Doo-Sabin subdivision seems to be a competitive, if not superior alternative. On one hand, generalizing quadratic B-splines, the expected approximation order for Doo-Sabin is $O(h^3)$ anyway so that extraordinary points do not have a negative impact. On the other hand, as mentioned above, the advantage of Doo-Sabin and possible trivariate variants is the smaller support of basis functions, implying sparser system matrices.

5 Integration

The challenges associated with numerical integration on subdivision surfaces are notorious. For bivariate Catmull-Clark subdivision, already [HKRD93] discussed them when optimizing surface fairness, and it is stated for instance in [WZHS15] that further study is needed to better understand the influence of quadrature rules on the convergence behavior of subdivision-based simulations. We will point to more references later on in this section.

Let us briefly sketch the issue. For the sake of simplicity, consider the biharmonic equation $\Delta^2 u = f$ with homogeneous Dirichlet boundary conditions on some domain $\Omega \subset \mathbb{R}^2$ as a prime example of an elliptic fourth order partial differential equation. If $\Omega$ is parameterized by the subdivision function $\varphi : S \to \Omega$, the Ritz-Galerkin approximation $u_h = z \circ \varphi^{-1}$ can be determined as follows: Let $\beta_k := b \circ \varphi^{-1}, k \in K$, denote the reparametrized subdivision basis, assumed to satisfy the boundary conditions, and
let $z = \sum b_i z_i$ be the unknown function. Then $a_{ij} = \sum \beta_i z_i$, and the vector $Z = (a_i)_i$ of coefficients is determined by the linear system $AZ = F$, where the entries $a_{ij}$ of the matrix and $f_j$ of the right-hand side are given by

$$a_{ij} = \int_\Omega \Delta \beta_i \Delta \beta_j \quad \text{and} \quad f_j = \int_\Omega f \beta_j,$$

For computational purposes, these integrals must be transferred from the physical domain $\Omega$ to the parametric domain $S$. We obtain

$$a_{ij} = \int_S \Delta \varphi_i \Delta \varphi_j \quad \text{and} \quad f_j = \int_S f \varphi^{-1} \beta_j,$$

where $\Delta \varphi_i$ denotes the Laplace-Beltrami operator with respect to $\varphi$, and $g := |\det D\varphi|$ is the measure factor. The following task concerning compilation of the linear system $AZ = F$ should not be underestimated.

**Task 5.1.** Develop efficient and highly precise numerical quadrature schemes for determining $A$ and $F$.

Remarkably, it is possible to derive exact values for integrals of the form $\int_\Omega \partial_i^\alpha b \partial_\beta b$ directly from the coefficients of the refinement equation. One can use either spectral methods [HR16] or the summation of certain geometric series [BNK18].

The presence of the parametrization $\varphi$ in the Laplace-Beltrami operator, however, prevents this simplicity for $a_{ij}$. Exact results for arbitrary $f_j$ cannot be expected anyway. Hence, there is no way around numerical integration.

The domain $S = \Sigma \times I$ consists of finitely many copies of the set $\Sigma \subset \mathbb{R}^d$. If, for instance, $\Sigma = [0, 1]^2$ is the unit square, the unit cube, it seems plausible to use standard quadrature schemes like tensor product Gauss-Legendre formulas for each such copy and to sum up the results, see e.g. [NKP14]. [JMPR16] refers to this approach as AG quadrature when comparing bivariate quadrature. The trial implementation discussed in Section 7 applies an alternative approach, similar to the GA approach in [HR16], using a fixed budget of points per mesh cell by locally subdividing until each Gauss point is enclosed by a regular neighborhood, so that we can evaluate.

However, typically, the optimal approximation power of numerical quadrature schemes is achieved only if the integrand is sufficiently smooth. For instance, Gauss-Legendre with $n^2$ nodes yields approximation order $O(n^{2d})$ only if the integrand is in $C^{2d}$. Discontinuities in lower order derivatives inevitably lead to a significant loss of accuracy. Dealing for instance with cubic splines, the integrand of $a_{ij}$ consists of smooth pieces, but, involving second order partials, neighboring pieces join merely continuously. Hence, it is sensible to treat each piece separately. In the subdivision setup, the challenge arises from the fact that the integrand in question is segmented into infinitely many smooth pieces. For instance, in the trivariate case, there is an infinite sequence of ever smaller hexahedral pieces accumulating near each extraordinary vertex.

The obvious approach to dealing with the situation is series truncation. That is, only the finitely many pieces whose contribution to the final integral is above a given threshold are treated, and the rest is simply ignored. Controlling this process is crucial and requires understanding the asymptotic behavior of the sub-integrals that are of the form

$$a_{ij}^{(\ell)} = \int_{S^\ell} \Delta \varphi_i \Delta \varphi_j \varphi^{-1} g^{(\ell)},$$

where $S^\ell$ is a cube in $\mathbb{R}^d$ with side length $2^{-\ell}$, arising as subdomain after $\ell$ steps of refinement, indexed by $\ell$. The superscripts of the involved functions indicate restriction to $S^\ell$. In $d$ dimensions, the volume of $S^\ell$ is $2^{-\ell d}$, the measure factor is $g^{(\ell)} = O((2\lambda)^{d^2})$, and $\Delta \varphi_i \Delta \varphi_j = O((\mu/\lambda)^d)$. Together, the asymptotic behavior is

$$a_{ij}^{(\ell)} = O((2\lambda)^{d^2} \lambda^{-d^2}).$$

Given that $\lambda \approx \lambda_0$ is fixed, the decay is characterized by the value of the subsubdominant eigenvalue $\mu$. The smaller that value, the more pieces fall below the threshold, the fewer pieces need to be integrated, and the shorter is the run time. Still, the number of integrals may be quite high so that optimal efficiency is mandatory for an implementation. A systematic treatment of the quadrature issue can be found in [JMPR16], where different approaches are described and compared numerically for the special case of Loop subdivision. Further, the methods developed in [BNK18] deserve special attention. In [BNK18] integration rules custom-tailored for Catmull-Clark surfaces are suggested for speedup. Yielding promising results, the approach should be generalized to the trivariate case. Also the approach developed in [Die18] is worth considering. Here, the infinitely many pieces beyond a given level are replaced by a finite approximating surface that is then used to perform the integration.

As a simplified example of adaptive numerical integration, Figure 3 shows a Catmull-Clark surface, whose area is computed using an adaptive algorithm and, for visualization purposes, the midpoint rule of integration. Parts near the extraordinary vertices are segmented into smaller pieces. The obvious approach is series truncation, but still above machine precision, the domain is split into smaller pieces. Estimates for the error resulting from truncation must be developed. One possible approach relies on the asymptotic expansion according to (2).
6 Linear independence and condition

As a last topic, we consider the Ritz-Galerkin system $AZ = F$ from a numerical point of view. First of all, it must be noted that it is not completely clear if the matrix $A$ has full rank or, to put it differently, if the basic limit functions of a subdivision algorithm are linearly independent. Even in the bivariate case, counterexamples are known, like the standard Catmull-Clark algorithm for an initial mesh with the connectivity of the edges of a cube [PW06]. Under refinement, such peculiarities typically disappear, but still, especially with regard to the trivariate case, one should make that rigorous.

**Task 6.1.** Show that—or provide conditions under which—the basic limit functions of a subdivision algorithm are linearly independent.

Of course, linear independence is a necessary prerequisite, but it is by no means sufficient. It is important to consider the condition of the matrix $A$, which may be regarded as a quantitative version of the notion of linear independence. Also here, the field is wide open.

**Task 6.2.** Derive at least qualitative bounds on the condition number of the Ritz-Galerkin matrix for second and fourth order elliptic partial differential equations in the subdivision setting.

When using bases with compact support, the typical asymptotic behavior of the condition number is $O(h^{-2})$ for second order and $O(h^{-4})$ for fourth order problems regardless of the specific choice of elements. Investigations should address the condition of adaptively refined functions spaces.

Modest condition is essential for the accuracy of numerical solutions and also for the rate of convergence of iterative solvers. In that respect, also strategies for preconditioning could be of interest, but this aspect should not be addressed before the many other, much more pressing problems discussed above are solved.

7 Numerical Experiments

A good number of investigations has explored numerical convergence of subdivision-based simulation. The range of outcomes, each based on a set of examples, partly motivates this paper. For Poisson’s equation and the $L^2$ norm, [PXXZ16] observe an approximation rate of $O(h^2)$ for functions on a flat bivariate domain. [MC17] numerically observe a max-norm convergence of $O(h^2)$ for extraordinary (valence $n$) vertices and an $L^2$ convergence of $O(h^{2.7})$. [ZSC18] state an $O(h^2)$ convergence for valence $n$ vertices with lower convergence rates as $n$ increases, and set out to improve the convergence constant. Similarly, [LWZ19] observe a convergence of $O(h^{2.3})$ deteriorating with increased $n$ and propose a new hybrid subdivision that converges $O(h^3)$ in...
numerical experiments. For surface fitting, which they view as an instance of IGA, [MM19] claim an improved rate of \(O(h^4)\) after adjusting the subsubdominant eigenspectrum. [WLZH21] use a nonuniform parametrization to observe 'optimal convergence rates' \(O(h^7)\) for Poisson's equation, but only suboptimal \(O(h^5)\) when solving the biharmonic equation. [PRY21] also observe \(O(h^5)\) convergence for 4th order error. [BnK18] attribute reduced convergence for Catmull-Clark splines to unboundedness of the derivatives of the second and third eigenfunctions for valencies \(n > 4\).

Preceding the branding of the approach as IGA, [COS00] present the algorithmic framework to non-tensor-product layout. (Throughout the 1980s and 1990s, the engineering community applied IGA to tensor-product spline surfaces under the name of higher-order isoparametric methods). Specifically, [COS00] apply Loop's subdivision to solving an obstacle course of well-known thin shell benchmark problems. The authors observe an energy error of \(O(h^2)\) and rapid reduction of the displacement error at a central point. An approximation rate of \(O(h^5)\) is also observed in [PXXZ16, PRC18, PRXC19] and [JMPR16] observe this convergence rate, with reduced accuracy for higher nodal valence \(n\), for their implementation of Laplace-Beltrami. For the bi-Laplacian they note that the one-point barycentric quadrature fails to yield convergence, but alternative quadratures yield \(O(h^2)\). They attribute the moderate convergence rate to Loop’s subdivision not being able to reproduce cubics for valence \(n > 6\).

[kHYL22] report that a variant of Loop subdivision achieves \(O(h^3)\) at the cost of poorer shape and [KLCD16] claim 'much faster' adaptive convergence.

Trivariate Catmull-Clark subdivision for IGA was pioneered in [BHU10] and further used in [AESF21]. [LMS20] report an \(L^2\) convergence of \(O(h^{2.5})\) for the 35-mesh (one slice of the 35-extruded mesh). For an interpolatory version of this approach, [XXD20, Fig 7] report an initial decay in the \(L^2\) error of \(O(h^{1.5})\) with the error concentrated at irregularities – but this observed order reduces to approximately \(O(h^2)\) by the third refinement step. To develop more intuition, and because it is of interest in its own right, we implemented trivariate Catmull-Clark subdivision following [MJ96, BHU10, YP22], i.e., by repeatedly choosing new points as centroids of their neighbors. This generalization of bivariate Catmull-Clark subdivision (see [BSWX01] for a competing version) is expected to be \(C^1\) at extraordinary points, \(C^2\) across extraordinary edges, and \(C^2\) elsewhere.

We tested convergence of the IGA formulation of both Poisson’s and the biharmonic equation on three box-complexes shown in Fig. 4. To measure convergence for Poisson's equation with zero boundary values, we chose

\[
f := (6\pi^2/9)\sin((\pi x)/3)\sin((\pi y)/3)\sin((\pi z)/3) \quad \text{on} \quad \Omega := [0, 6]^3. \tag{3}
\]

Then the exact solution of Poisson’s equation is

\[
u = 2\sin\left(\frac{\pi x}{3}\right)\sin\left(\frac{\pi y}{3}\right)\sin\left(\frac{\pi z}{3}\right). \tag{4}\]

We refine each of the three types of meshes, solve and measure the errors. The number of tensor grid boxes for refinement \(0, 1, 2, 3, 4\) are 8, 64, 512, 4,096, and 32,768. The counts for 35-extruded mesh and sphere octant are similar. The bulk of the time is therefore spent on assembling the stiffness matrix \(A\).

We subdivided locally until each Gauss point is enclosed by a regular neighborhood, so that we can evaluate a tricubic polynomial. This approach uses a fixed budget, one of the options considered in [JMPR16], as opposed to [NKP14] where subdivision rings are generated and integrated until the error is sufficiently small. We compared integration at \(k^3\) Gauss points where \(k \in \{3, 5, 7, 9\}\) and found that the \(L^2\) error (and similarly \(L^\infty\) and \(H^1\) errors) for 3, respectively 4, refinement steps differed by less than 0.5% for the range of \(k\). That is, any lack of convergence is unlikely to be due to insufficient sampling for integration at Gauss points.

We tested convergence of the IGA formulation of both Poisson’s and the biharmonic equation on three box-complexes shown in Fig. 4. To measure convergence for Poisson’s equation with zero boundary values, we chose

\[
f := (6\pi^2/9)\sin((\pi x)/3)\sin((\pi y)/3)\sin((\pi z)/3) \quad \text{on} \quad \Omega := [0, 6]^3. \tag{3}
\]

Then the exact solution of Poisson’s equation is

\[
u = 2\sin\left(\frac{\pi x}{3}\right)\sin\left(\frac{\pi y}{3}\right)\sin\left(\frac{\pi z}{3}\right). \tag{4}\]

We refine each of the three types of meshes, solve and measure the errors. The number of tensor grid boxes for refinement \(0, 1, 2, 3, 4\) are 8, 64, 512, 4,096, and 32,768. The counts for 35-extruded mesh and sphere octant are similar. The bulk of the time is therefore spent on assembling the stiffness matrix \(A\).

We subdivided locally until each Gauss point is enclosed by a regular neighborhood, so that we can evaluate a tricubic polynomial. This approach uses a fixed budget, one of the options considered in [JMPR16], as opposed to [NKP14] where subdivision rings are generated and integrated until the error is sufficiently small. We compared integration at \(k^3\) Gauss points where \(k \in \{3, 5, 7, 9\}\) and found that the \(L^2\) error (and similarly \(L^\infty\) and \(H^1\) errors) for 3, respectively 4, refinement steps differed by less than 0.5% for the range of \(k\). That is, any lack of convergence is unlikely to be due to insufficient sampling for integration at Gauss points.

We tested convergence of the IGA formulation of both Poisson’s and the biharmonic equation on three box-complexes shown in Fig. 4. To measure convergence for Poisson’s equation with zero boundary values, we chose

\[
f := (6\pi^2/9)\sin((\pi x)/3)\sin((\pi y)/3)\sin((\pi z)/3) \quad \text{on} \quad \Omega := [0, 6]^3. \tag{3}
\]

Then the exact solution of Poisson’s equation is

\[
u = 2\sin\left(\frac{\pi x}{3}\right)\sin\left(\frac{\pi y}{3}\right)\sin\left(\frac{\pi z}{3}\right). \tag{4}\]

We refine each of the three types of meshes, solve and measure the errors. The number of tensor grid boxes for refinement \(0, 1, 2, 3, 4\) are 8, 64, 512, 4,096, and 32,768. The counts for 35-extruded mesh and sphere octant are similar. The bulk of the time is therefore spent on assembling the stiffness matrix \(A\).

We subdivided locally until each Gauss point is enclosed by a regular neighborhood, so that we can evaluate a tricubic polynomial. This approach uses a fixed budget, one of the options considered in [JMPR16], as opposed to [NKP14] where subdivision rings are generated and integrated until the error is sufficiently small. We compared integration at \(k^3\) Gauss points where \(k \in \{3, 5, 7, 9\}\) and found that the \(L^2\) error (and similarly \(L^\infty\) and \(H^1\) errors) for 3, respectively 4, refinement steps differed by less than 0.5% for the range of \(k\). That is, any lack of convergence is unlikely to be due to insufficient sampling for integration at Gauss points.

We subdivided locally until each Gauss point is enclosed by a regular neighborhood, so that we can evaluate a tricubic polynomial. This approach uses a fixed budget, one of the options considered in [JMPR16], as opposed to [NKP14] where subdivision rings are generated and integrated until the error is sufficiently small. We compared integration at \(k^3\) Gauss points where \(k \in \{3, 5, 7, 9\}\) and found that the \(L^2\) error (and similarly \(L^\infty\) and \(H^1\) errors) for 3, respectively 4, refinement steps differed by less than 0.5% for the range of \(k\). That is, any lack of convergence is unlikely to be due to insufficient sampling for integration at Gauss points.

We subdivided locally until each Gauss point is enclosed by a regular neighborhood, so that we can evaluate a tricubic polynomial. This approach uses a fixed budget, one of the options considered in [JMPR16], as opposed to [NKP14] where subdivision rings are generated and integrated until the error is sufficiently small. We compared integration at \(k^3\) Gauss points where \(k \in \{3, 5, 7, 9\}\) and found that the \(L^2\) error (and similarly \(L^\infty\) and \(H^1\) errors) for 3, respectively 4, refinement steps differed by less than 0.5% for the range of \(k\). That is, any lack of convergence is unlikely to be due to insufficient sampling for integration at Gauss points.

The coefficients of the outer boundary of the mesh were set to zero to enforce (zero) Dirichlet boundary conditions. We used Matlab’s preconditioned conjugate gradient solver with the settings

\[
L = 
ichol(A, struct('type','ict','droptol',1e-3,'diagcomp',10.0));
\]

\[
x = pcg(A, b, [], 100000, L, L');
\]

The solver always achieved \(|AZ - F|/|F| < 10^{-6}\) and the output agreed to relevant accuracy with that of direct solvers, for refinement levels 1 and 2. Since, without refinement the control nets do not offer sufficient degrees of freedom for a sensible solution and the top entry of Table 1 and Table 2 are therefore empty.

![tensor grid](image1)

![35-extruded mesh](image2)

![sphere octant](image3)

**Figure 5:** Volumetric Catmull-Clark convergence of the error for Poisson’s equation on the tensor grid, 35-extruded mesh, sphere octant box-complexes.

The graphs in Fig. 5 display the three error ratios of the four measurements with ‘refs’ refinements. The error ratios for Fig. 5a are, as expected, consistent with optimal \(O(h^4)\) convergence. However Fig. 5b,c show the convergence to the solution of Poisson’s equation to be, as predicted, approximately \(O(h^2)\).
We believe that subdivision can provide a powerful framework for the simulation of higher order PDEs in the trivariate case. Quadrature with 4

\[ \text{Jörg Peters and Jeremy Youngquist were supported in part by DARPA TRADES HR00111720031.} \]

\subsection*{Acknowledgements}
Jörg Peters and Jeremy Youngquist were supported in part by DARPA TRADES HR00111720031.

\section*{8 Conclusion}
We believe that subdivision can provide a powerful framework for the simulation of higher order PDEs in the trivariate case. However, many tasks have to be completed to place computations on solid ground. The tasks include the development of appropriate algorithms, their regularity analysis, the development of efficient quadrature schemes, and a better understanding of approximation and stability properties. Some of these tasks may prove to be challenging. Ideally, joint efforts of the subdivision community will advance understanding and help to achieve the goal we have in mind.
35-extruded mesh

sphere octant

Figure 6: Error decay for the biharmonic equation solved by IGA leveraging Catmull-Clark solids.

References


