Journal of Computational Physics 281 (2015) 352–364

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

# A mesh partitioning algorithm for preserving spatial locality in arbitrary geometries

Girish V. Nivarti [*,1], M. Mahdi Salehi, W. Kendal Bushe

*Department of Mechanical Engineering, University of British Columbia, 6250 Applied Science Lane, Vancouver, British Columbia, V6T 1Z4, Canada*

A R T I C L E I N F O

*Article history:*
Received 26 November 2013
Received in revised form 30 June 2014
Accepted 10 October 2014
Available online 24 October 2014

*Keywords:*
Mesh partitioning
Space-filling curves
Computational fluid dynamics
Parallel computing
Turbulent reactive flows

A B S T R A C T

A space-filling curve (SFC) is a proximity preserving linear mapping of any multi-dimensional space and is widely used as a clustering tool. Equi-sized partitioning of an SFC ignores the loss in clustering quality that occurs due to inaccuracies in the mapping. Often, this results in poor locality within partitions, especially for the conceptually simple, Morton order curves. We present a heuristic that improves partition locality in arbitrary geometries by slicing a Morton order curve at points where spatial locality is sacrificed. In addition, we develop algorithms that evenly distribute points to the extent possible while maintaining spatial locality. A metric is defined to estimate relative inter-partition contact as an indicator of communication in parallel computing architectures. Domain partitioning tests have been conducted on geometries relevant to turbulent reactive flow simulations. The results obtained highlight the performance of our method as an unsupervised and computationally inexpensive domain partitioning tool.



## 1. Introduction

Data classification is a frequently encountered solution approach in a wide range of scientific disciplines. Spatial partitioning is a classification technique that is critical for obtaining accurate solutions of physical problems wherein constitutive local phenomena demand segregated treatment. In complicated domain geometries, solutions can be achieved conveniently by partitioning the problem space into localised sub-domains. Moreover, with the advent of parallel processing technology, domain partitioning has become an essential tool to address large-scale problems by dividing computational load across the available resources. In this context, size governs load balance whereas locality designates communication between processors. The choice between load and locality is particularly difficult to make for problems with multiple constraints.

The partitioning technique presented in the current work is motivated by a stringent demand posed by a relatively recent model, termed Conditional Source-term Estimation (CSE), applied for the computational fluid dynamics (CFD) simulation of turbulent reacting flows [1–4]. Conditionally averaged scalars, which are required for the CFD solution, are assumed to be statistically homogeneous in spatially localised ensembles of points. Within a localised ensemble, an integral equation may be accurately inverted only if a certain threshold number of computationally significant points are available. These

* Corresponding author.
*E-mail address:* g.nivarti@alumni.ubc.ca (G.V. Nivarti).
[1] Current address: Department of Engineering, University of Cambridge, Trumpington Street, Cambridge, CB2 1PZ, United Kingdom. Tel.: +44 1223 748 577.

http://dx.doi.org/10.1016/j.jcp.2014.10.022
0021-9991/

constraints pose stringent demands for both computational load balance and spatial locality of the ensemble generation or mesh partitioning strategy adopted.

In simple geometries, equi-sized slices of the domain along its major axis have sufficed [5,6]. In contrast, automated construction of ensembles is absolutely necessary for applying the model to the arbitrarily complex geometries that have industrial relevance. But, spatial proximity *must* be maintained within each cluster to retain the validity of the underlying assumption. In addition, parallel implementation of this large-scale problem demands that a measure of communicational load balance be considered in the partitioning procedure. Precise estimation of element-wise computational loads and the subsequent implementation of a multi-constraint partitioning procedure has not been addressed in this work as this requires, among other programming factors, an *a priori* knowledge of the critical fields. Finally, frequent mesh changes enforced by industrial demands or by adaptive mesh refinement (AMR) impose a limit on the computational expense of the partitioning algorithm.

What makes this problem particularly challenging is that locality is not merely desirable from a standpoint of computational efficiency, but is critical for calculating a valid solution. Load balance has relatively less significance, and as such, were there to be a trade-off between load balance and strict locality, strict locality must be preferred. In fact, a trade-off is *necessary* because strict locality alone does not guarantee a solution for the governing integral equation; in the absence of *a priori* knowledge of critical fields, a threshold ensemble size must be maintained.

### *1.1. Relevant background*

The problem of clustering $N$ objects into $M$ groups of objects with similar properties is a persistent challenge in several academic fields. Frequent applications are found in computer vision, pattern recognition, geographical information systems, large-scale database management, cache performance and many other fields [7,8]. Statistical classification, such as the cognitive mapping developed by Jenks [9], seeks to reduce the variance within classes and maximise the variance between classes. Hartigan [10] discusses the relevant statistical theory for clustering, whereas Davé and Krishnapuram [11] present a unified view of robust clustering methods. These age old methods are efficient for applications with few dimensions, but it is recognised that the *curse of high-dimensionality* pervades the problem of clustering nearest point neighbours [12].

In CFD problems, partitioning corresponds to clustering regions of mesh based elements such as nodes, finite elements or finite control volumes (also called *cells*). For a multi-block finite volume mesh, popularly used in AMR implementations, blocks of cells can be divided into clusters assigned to different processors in the computing architecture. Such computational mesh partitioning has been the focus of extensive research [13–16]. The choice of adopted strategy depends on known geometric information and the trade-off between partition quality and computational expense that best suits the problem. Dynamic problems are addressed using fast methods, whereas relatively slower but highly customised procedures are formulated to match specific criteria such as those imposed by the combustion model.

Over the last few decades, a variety of mesh partitioning techniques have been developed, the majority of which fall into two broad classes. Geometric algorithms including recursive bisection techniques and the SFC approach need only the local geometric information, such as block-centre coordinates. As an example, the recursive coordinate bisection (RCB) developed by Berger and Bokhari [17] is widely accepted due to its simplicity and the wide range of its applicability. The unbalanced recursive bisection strategy provides an advantageous modification of the RCB method [18] and Gilbert et al. [19] investigate yet another geometric method of dividing an irregular mesh into equal-sized pieces with few interconnecting faces. In general, geometric partitioning is conceptually simple, so the corresponding algorithms are fast and easy to implement. Additionally, the SFC approach, discussed at greater length in Section 1.2, provides a unified and scalable data structure relevant for a wide range of CFD operations [20,21]. Several software distributions such as the Zoltan toolkit [22] are available today that package a variety of partitioning schemes together.

The other broad class of partitioning, graph partitioning, has been reviewed by Schloegel et al. [23]. In this method, nodes of a computational graph represent tasks that can be executed concurrently and edges represent the communication required between tasks from one iteration to the next. Domain partitioning using such graphs is challenging and is recognised as an NP-hard problem [24]. The greedy algorithm developed by Farhat and Lesoinne [25], *builds* each partition by starting with a vertex and adding adjacent vertices until the target size or expected computational load has been reached. A relatively advanced technique, termed recursive spectral bisection (RSB), was developed by Simon [13]. It provides excellent quality partitions at exceedingly high computational costs. Many standard graph-based partitioning packages have been developed to employ efficient global and local graph-based methods for domain partitioning. Karypis and Kumar [26] have developed a scalable parallel algorithm for irregular graph partitioning that is applicable to finite element/volume meshes. Recently, Teresco et al. [27] provided an comprehensive review of partitioning algorithms with an extensive section on graph-based methods.

As mentioned earlier, the choice of adopted strategy depends on known information and the expense-quality trade-off that best suits the problem. The SFC approach not only provides a cheap clustering technique, but also an efficient data structure for performing CFD operations faster. In this work, we employ an SFC to partition a multi-block finite volume mesh.