

# From $h$ to $p$ efficiently: selecting the optimal spectral/ $hp$ discretisation in three dimensions

C. D. Cantwell<sup>a1</sup>, S. J. Sherwin<sup>b</sup>, R. M. Kirby<sup>c</sup> and P. H. J. Kelly<sup>d</sup>

<sup>a</sup> Department of Mathematics, Imperial College London, London, SW7 2AZ, UK

<sup>b</sup> Department of Aeronautics, Imperial College London, London, SW7 2AZ, UK

<sup>c</sup> School of Computing, University of Utah, Salt Lake City, Utah, USA

<sup>d</sup> Department of Computing, Imperial College London, London, SW7 2AZ, UK

**Abstract.** There is a growing interest in high-order finite and spectral/ $hp$  element methods using Continuous and Discontinuous Galerkin formulations. In this paper we investigate the effect of  $h$ - and  $P$ -type refinement on the relationship between runtime performance and solution accuracy. The broad spectrum of possible domain discretisations makes establishing a performance-optimal selection non-trivial. Through comparing the runtime of the optimal operator implementation at each discretisation with a desired solution tolerance we demonstrate how the optimal discretisation may be selected for a specified problem. Furthermore, this demonstrates the essential need for codes to support both low- and high-order discretisations.

**Key words:** spectral/ $hp$  element, optimisation, code performance

**AMS subject classification:**

## 1 Introduction

The use of low- and high-order finite element methods has become a mainstream technique in the solution of partial differential equations (PDEs). The ability to handle complex geometries, provide localised refinement in regions of high gradient and attain convergence at exponential rates make spectral/ $hp$  element methods valuable in a wide variety of applications across academia and industry. These originally included problems in fluid dynamics [11], but are now used in many

---

<sup>1</sup>Corresponding author. E-mail: c.cantwell@imperial.ac.uk

other fields such as studies of electromagnetics [6], shallow-water problems [1] and structural dynamics [9].

Spectral/ $hp$  element methods can be considered as a high-order extension of the traditional low-order finite element methods. They present greater scope for achieving convergence of a solution through both refining the mesh ( $h$ -refinement) and increasing the order of the polynomial expansions ( $P$ -refinement) used to represent the solution. Global spectral techniques can be considered as the limit of  $P$ -refinement on a single element. The notion of high-order is understood differently depending on the context. Those practising  $h$ -type refinement consider high-order expansions to be anything up to fifth-order [7, 16] while at the other end of the spectrum, the global spectral community would consider expansion orders in the region of 100 to be relatively modest [5]. In contrast, the spectral/ $hp$  element community [8, 3] would place high-order in the region of 15th-order.

Spectral/ $hp$  discretisations offer flexibility not only through mesh refinement and expansion order but through selection of one of several evaluation strategies. Changing the implementation strategy used with a given discretisation affects solver performance considerably, particularly in three dimensions, as discussed in Cantwell *et al.* [2]. The choice of discretisation controls the solution accuracy. In solving a given PDE using a spectral/ $hp$  element framework, one typically strives to achieve a required level of accuracy at minimal runtime cost. Therefore one must find the optimal  $(h, P)$  discretisation which achieves this. In practice, other factors may influence the choice such as memory limitations, but we discount these for the purposes of this paper. Cantwell *et al.* examined the runtime performance of a spectral/ $hp$  implementation in actioning various fundamental operators (including transforms and linear differential operators such as the Helmholtz operators) in three dimensions. They found that choosing the correct implementation strategy when evaluating the different operators was essential to ensure an efficient implementation through minimising runtime performance for a given  $(h, P)$  discretisation.

In this paper we establish the route to solving an elliptic problem, the Helmholtz equation, to a given accuracy threshold with minimal runtime cost through variation of these discretisation parameters using the optimal evaluation strategy for each  $(h, P)$ -combination as established in [2]. We examine how an optimal discretisation might be selected for a given problem by establishing the choice of  $h$  and  $P$ , and consequently the optimal evaluation strategy, which minimises the runtime while computing the solution to within a prescribed error tolerance. Furthermore, the most efficient type of expansion (hexahedral or tetrahedral) for the problem is established.

The paper is organised as follows. In Sec. 2 we summarise the spectral/ $hp$  element formulation and describe the three operator implementation strategies which may be adopted for a particular discretisation. We also describe the example problem used to demonstrate the optimal discretisation selection. In Sec. 3 we investigate how the optimal combination of mesh size and expansion order may be selected to solve the example problem to a prescribed level of accuracy in minimal runtime. Finally, we discuss the consequences and limitations of the work in Sec. 4.

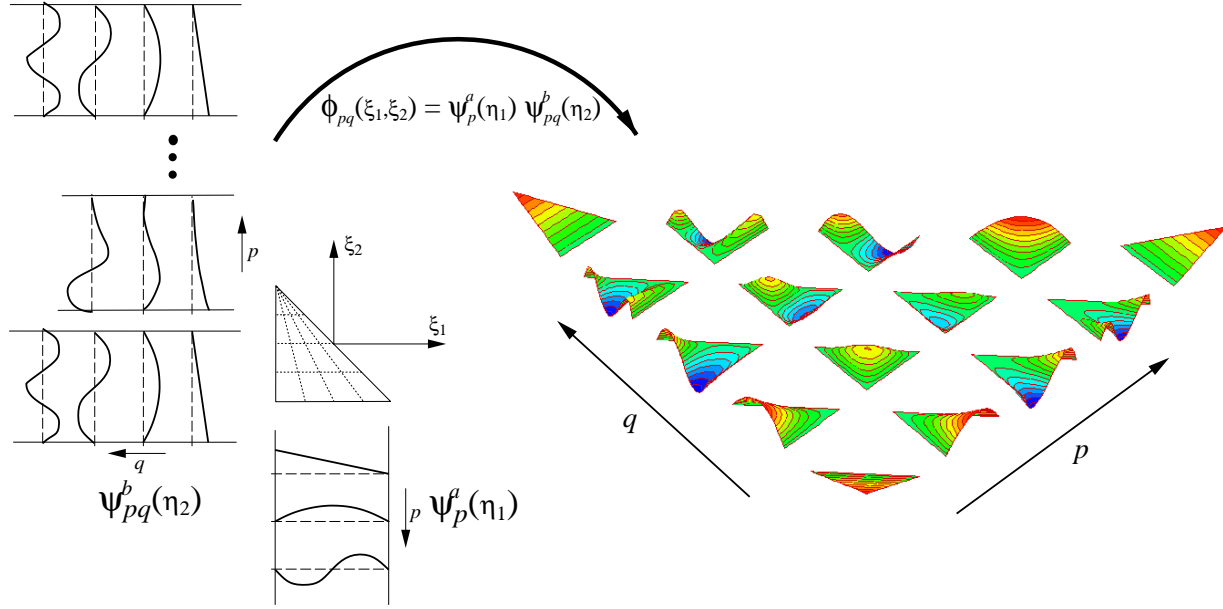


Figure 1: Two-dimensional triangular modes  $\phi_{pq}(\xi_1, \xi_2)$  are constructed as a tensor product of one-dimensional modes  $\psi_p(\eta_1)$  and  $\psi_{pq}(\eta_2)$ .

## 2 Methods

### 2.1 Spectral/hp element framework

The spectral/hp element framework and the tensorial construction of hexahedral and tetrahedral elements is well-established and is documented extensively elsewhere ([8, 11, 12, 13]). Therefore, we only give a brief summary of the formulation here.

A potentially complex domain is partitioned into a tessellation of elemental sub-domains. In the three-dimensional setting we consider an element  $\Omega_e$  to be either hexahedral or tetrahedral in shape, although prisms and square-based pyramids may also be used. For each  $\Omega_e$ , a mapping  $\chi_e : \Omega_e \rightarrow \Omega_{st}$  maps points in the physical elemental subdomains onto the corresponding shape in reference space,  $\Omega_{st}$ . The reference-space hexahedral and tetrahedral regions are defined as

$$\begin{aligned} \mathcal{Q}^3(\xi) &= \{(\xi_1, \xi_2, \xi_3) \in [-1, 1]^3\}, \\ \mathcal{T}^3(\xi) &= \{(\xi_1, \xi_2, \xi_3) \mid -1 \leq \xi_i; \xi_1 + \xi_2 + \xi_3 \leq -1\}, \end{aligned}$$

respectively. For the latter, a coordinate transform [4, 12] leads to a representation of the tetrahedral geometry with fixed limits as

$$\mathcal{T}^3(\eta) = \{(\eta_1, \eta_2, \eta_3) \in [-1, 1]^3\}.$$

On each reference element, the solution for a given element is expressed in terms of a fixed basis of three-dimensional functions,  $\{\phi_n\}$ , constructed as a tensor product of  $P$  one-dimensional

polynomials  $\{\psi_p\}$ . This leads to the following expansions on the reference elements

$$\mathcal{Q}^3 : \quad u(\xi) = \sum_n \phi_n \hat{u}_n = \sum_p \sum_q \sum_r \psi_p(\xi_1) \psi_q(\xi_2) \psi_r(\xi_3) \hat{u}_{pqr} \quad (2.1)$$

$$\mathcal{T}^3 : \quad u(\eta) = \sum_n \phi_n \hat{u}_n = \sum_p \sum_q \sum_r \psi_p(\eta_1) \psi_{pq}(\eta_2) \psi_{pqr}(\eta_3) \hat{u}_{pqr} \quad (2.2)$$

To ensure only the minimum degrees of freedom necessary are used in the tetrahedral case (an extension to three dimensions of the triangular case shown in Fig. 1) there is an interdependency of one-dimensional modes in the second and third coordinate directions. The nature of this construction allows for the sum-factorisation evaluation strategy [10] to be employed, even in the tetrahedral region, which leads to a higher performance in some circumstances [2].

In a continuous Galerkin spectral/ $hp$  element formulation the  $C^0$ -continuity requirement is imposed across elemental boundaries to enforce the required connectivity. This leads to a global formulation of the problem in which each of the elemental modes are considered as global functions and assigned a global numbering. An assembly matrix  $\mathcal{A}$  maps this global mode numbering onto the corresponding local elemental mode numberings,

$$\hat{\mathbf{u}}^l = \mathcal{A} \hat{\mathbf{u}}^g.$$

This is simplified if the polynomial functions used in the construction of the basis support a boundary/interior decomposition. In this case all elemental interior modes are zero on the element boundaries, and consequently interior modes on different elements are globally orthogonal. This leads to a global matrix system that exposes significant structure which may be exploited using substructuring techniques [14] to improve the performance of global operations.

We conclude the formulation with an overview of the implementation strategies with which we may evaluate numerical operators. As shown in [2] for the three-dimensional case, and in [15] in two dimensions, careful selection of the implementation strategy is essential to achieve the best performance at a given  $(h, P)$ -discretisation. As well as global operations and local elemental operations the tensor-product construction of the elemental basis modes allows for a third method of implementation known as sum-factorisation [10]. Therefore, the three methods we consider for evaluating operators are

- Global matrix evaluation,
- Local elemental matrix evaluation,
- Local elemental sum-factorisation.

A diagrammatic representation of these is given in Fig. 2 showing the global approach (Fig. 2(a)) and the two elemental approaches (Fig. 2(b) and (c)), as well as the elemental decomposition of the domain.

The global approach evaluates all degrees of freedom in the domain simultaneously through the use of a global matrix system. The local strategies perform the evaluation at the elemental level.

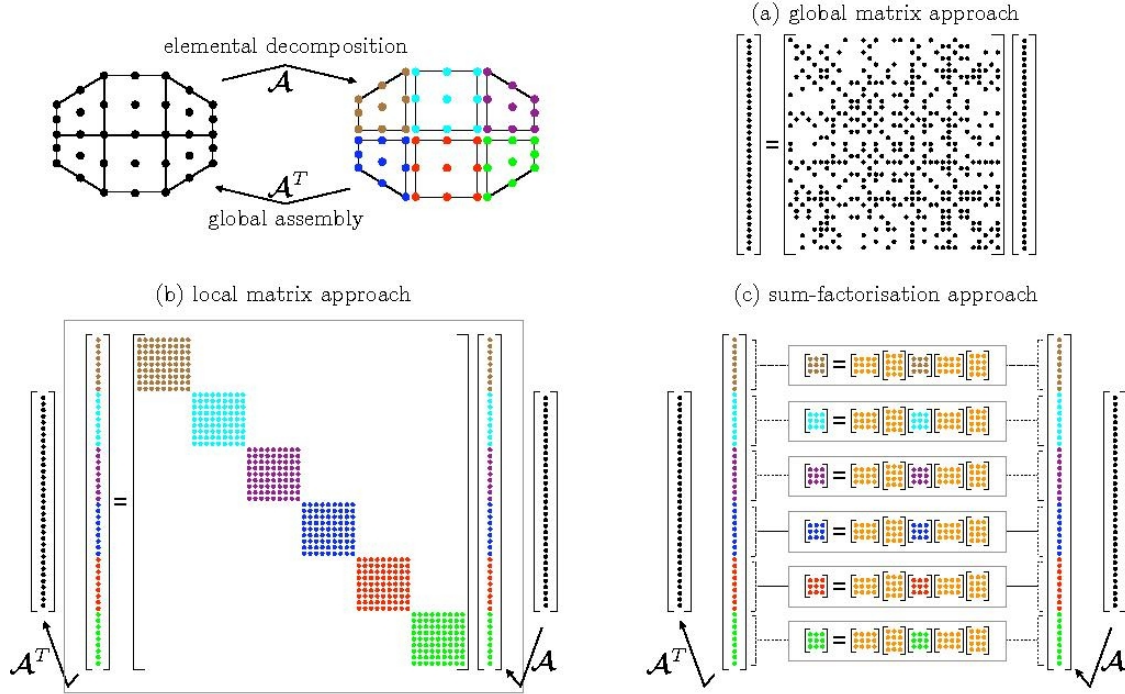


Figure 2: Diagram of the three evaluation strategies: global matrix operation (a), local elemental operations (b) and sum-factorisation (c).

The sum-factorisation approach factorises the fully three-dimensional  $\mathcal{O}(P^6)$  elemental operation as a series of three one-dimensional  $\mathcal{O}(P^4)$  operations. This provides significantly improved performance with some discretisations [2], reducing both memory storage and computational cost. This is not always the case, and with some linear differential operators with certain discretisations it may provide a significantly lower performance than the local matrix or global strategies. To illustrate the sum-factorisation implementation, consider the backward transform operation for the hexahedral region in Eqn. 2.2. This is factorised as

$$u(\xi_{1i}, \xi_{2j}, \xi_{3k}) = \sum_{p=0}^P \psi_p(\xi_{1i}) \left\{ \sum_{q=0}^P \psi_q(\xi_{2j}) \left\{ \sum_{r=0}^P \hat{u}_{pqr} \psi_r(\xi_{3k}) \right\} \right\}, \quad (2.3)$$

which may in turn be expressed as the matrix equation.

$$\mathbf{U}_{[P]} = \left[ \left( \hat{\mathbf{U}}_{[P]}^\top \mathbf{B}_0^\top \right)^\top \mathbf{B}_1^\top \right]^\top \mathbf{B}_2^\top = \mathbf{B}_1 (\hat{\mathbf{U}}_{[P]}^\top \mathbf{B}_0^\top) \mathbf{B}_2^\top.$$

In this form, the operation may be computed particularly efficiently in the hexahedral region, especially if the benefits of an optimised BLAS are available.

## 2.2 Example problem

The problem considered is that of solving the Helmholtz equation

$$\nabla^2 u - \lambda u = f, \quad (2.4)$$

on the unit cube  $[0, 1]^3$  with  $\lambda > 0$ . The domain is equipped with suitable Dirichlet boundary conditions appropriate to the forcing function  $f$ . This is a typical calculation encountered in a broad range of applications and therefore makes an effective example of the optimisation techniques which may be applied within a spectral/hp framework. We consider the forcing function

$$f = -(\lambda + 3n^2\pi^2) \sin(n\pi x) \sin(n\pi y) \sin(n\pi z). \quad (2.5)$$

which has exact smooth solution

$$u(x, y, z) = \sin(n\pi x) \sin(n\pi y) \sin(n\pi z). \quad (2.6)$$

The integer  $n$  parametrises the family of solutions, with high frequency forcing functions leading to larger solution errors. These are measured in the  $L_2$ -norm and defined as:

$$E = \left[ \int_{\Omega} (u_{exact} - u)^2 \right]^{\frac{1}{2}}.$$

The error is calculated using 30 quadrature points, rather than the  $P + 1$  points necessary to support an order  $P$  basis, to avoid poor measurement through aliasing effects with low-order expansions. The choice of  $n$  controls the solution frequency and high frequency solutions will be captured less accurately by a fixed discretisation. While for increasing  $n$  there is a quantitative difference in the convergence of the error as the discretisation is refined, the qualitative picture remains comparable. Therefore, for the results presented in the following section the choice of  $n = 3$  is used.

## 2.3 Test system

The dependence on hardware of the timings make the specification of the system particularly relevant. All calculations presented in this paper are conducted on the same MacPro 64-bit system with dual quad-core 2.26Ghz Xeon processors and 16GB system memory. The implementation of BLAS is the Accelerate framework included with OSX. All matrix operations (including global matrices) are applied using direct techniques, rather than iterative methods.

## 3 Optimal Strategy in three-dimensions

We consider the case of hexahedral elements first. Figure 3(a) summarises a comparison of the run-times of the different strategies for evaluating the Helmholtz operator using hexahedral elements at a fixed mesh density of  $h = 4$ . In summary, global strategies offer better performance at low

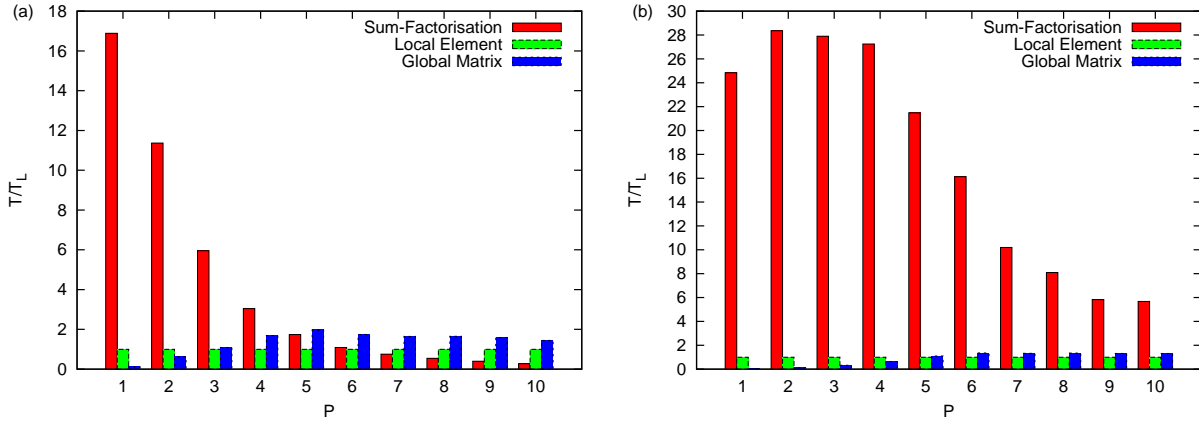


Figure 3: Comparison of evaluation strategy performance at different spectral/ $hp$  expansion orders for (a) hexahedral elements, (b) tetrahedral elements.

expansion orders, while sum-factorisation offers better performance in high-order regimes. Local elemental evaluation is the best approach for certain intermediate orders. We define the *optimal strategy* to be the strategy which minimises runtime for each  $(h, P)$ -pair. This will be the strategy used to select the optimal discretisation.

Figure 4 shows logarithmic contours of runtime (dotted lines) overlaid with logarithmic contours of  $L_2$ -error (solid lines). The first three plots (Fig. 4(a-c)) show contours of runtime for sum-factorisation, local elemental evaluation and global evaluation. Figure 4(d) shows runtime of the optimal strategy. Contours of error are the same on all plots in the figure. The apparent distortion of the  $10^{-3}$  error contour in the hexahedral case is an artifact of the  $(h, P) = (4, 3)$  case representing the  $n = 3$  solution particularly accurately due to the alignment of the  $h$ -mesh with our chosen solution. A similar artifact is apparent at  $(h, P) = (6, 3)$ .

We begin by briefly relating our present understanding of strategy performance from Fig. 3(a) and the results for the hexahedral region presented in Cantwell *et al.* [2] to the data in Fig. 4. The sum-factorisation strategy for the Helmholtz operator is observed to perform poorly at low orders. In Fig. 4 this is apparent from a quantitative comparison of the runtime contours across the first three plots. The global strategy performs best at low orders, again notable from the low contour values. The vertical orientation of the runtime contours at low expansion orders suggests that this strategy also scales well with increasing mesh density in this low-order regime. This is in contrast to the sum-factorisation contours which are predominantly horizontal and therefore favour scaling with polynomial order. A final point to note from the first three implementation plots is that for the local elemental and global strategies the runtime contours are parallel and with similar contour separation. This aligns with the observation from Fig. 3 that the relative increase in runtime of the sum-factorisation strategy with polynomial order reaches a plateau of approximately twice the runtime of the local elemental approach.

We next examine how the optimal discretisation may be selected for the three strategies separately. For the case of the sum-factorisation, the orientation of the contours of runtime compared with the contours of error lead to an optimal discretisation of a small number of high-order ele-

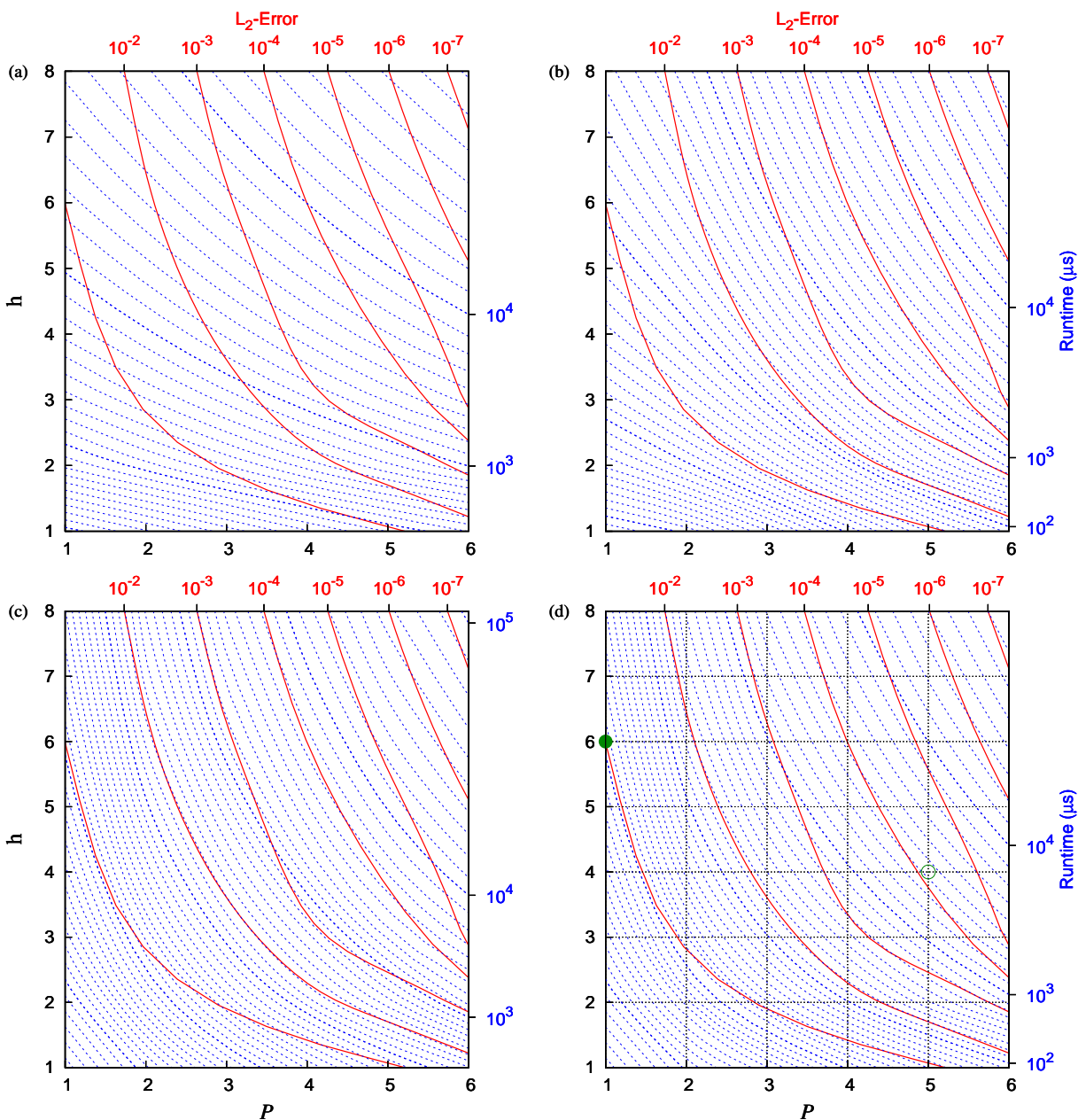


Figure 4: Contour plots showing the runtime (dotted lines) and  $L_2$ -error (solid lines and fixed across all plots) for each  $(h, P)$ -combination in solving the Helmholtz problem using hexahedral elements. The three evaluation strategies are shown: sum-factorisation (a), elemental matrices (b) and global matrix (c). A comparison with the optimal strategy chosen for each discretisation is shown in (d), where the filled circle marks the optimal discretisation to attain a solution with a 10% error tolerance, while the open circle indicates the optimal discretisation for 0.01%.

ments. This is true independent of the desired solution accuracy. In contrast, the global strategy is more complicated. The runtime and error contours are largely parallel leading to a broad spectrum of appropriate discretisations. However, for large error tolerances the marginally optimal choice is to favour mesh density over high order, while for small tolerances the exponential convergence dominates leading to high-order providing better performance. The local elemental approach is more interesting still. For a more relaxed error tolerance neither  $h$ -refinement or  $p$  refinement is dominant and a middle-ground discretisation is more appropriate. However, as with the global strategy, high-order dominates when accuracy is required.

The final sub-figure, Fig. 4(d), shows the comparison for the optimal strategy. This essentially merges the results from Fig. 4(a)-(c) to show the discretisation which offers the best possible performance. The character of the optimal strategy contours can be seen to contain aspects from the three base strategies, most notably from the global strategy at low orders and the local elemental and sum-factorisation strategies at high-orders. This aligns with the timing results of Fig. 3(a). Of course, in selecting an optimal discretisation only discrete selections of  $h$  and  $P$  are possible and the intersection of the grid lines denotes the available choices. In particular, we highlight the optimal discretisation for a 10% error tolerance (marked by the filled circle on the figure), and the optimal discretisation for a 0.01% tolerance (marked by the open circle). We will discuss these choices in the next section.

We now address the analysis of tetrahedral elements in Fig. 5. The format of the plots in this figure are identical to Fig. 4, showing error and runtime for the three base strategies and the optimal case. A comparison of the performance for the different strategies is shown in Fig. 3(b). The selection of optimal discretisation for these strategies is much more straightforward. For both sum-factorisation and local elemental strategies a high-order approach is favoured in all cases, while for the global approach a low-order finite element discretisation offers the best performance. The latter is most likely a consequence of the multiplicity of the degrees of freedom in a tetrahedral mesh.

Again we show the optimal strategy for the tetrahedral case and mark the optimal discretisations for a 10% and 0.1% error tolerance. For the optimal implementation the approach to take is dependent on the solution accuracy required, in much the same way as for the hexahedral case. With a high tolerance of error, one should select a low-order discretisation, while for greater accuracy high-order elements are essential for optimal performance. Finally, it can be noted that qualitatively hexahedral elements offer better performance over tetrahedral elements for this particular problem with a difference of half an order of magnitude in runtime for both the  $10^{-1}$  and  $10^{-4}$  error contours.

## 4 Discussion

The selection of an optimal discretisation for a problem can be appreciated to be a non-trivial measurement. It is dependent not only on the strategy employed by the spectral/ $hp$  solver, but on the nature of the problem and the desired accuracy. In this paper we have computed and analysed how

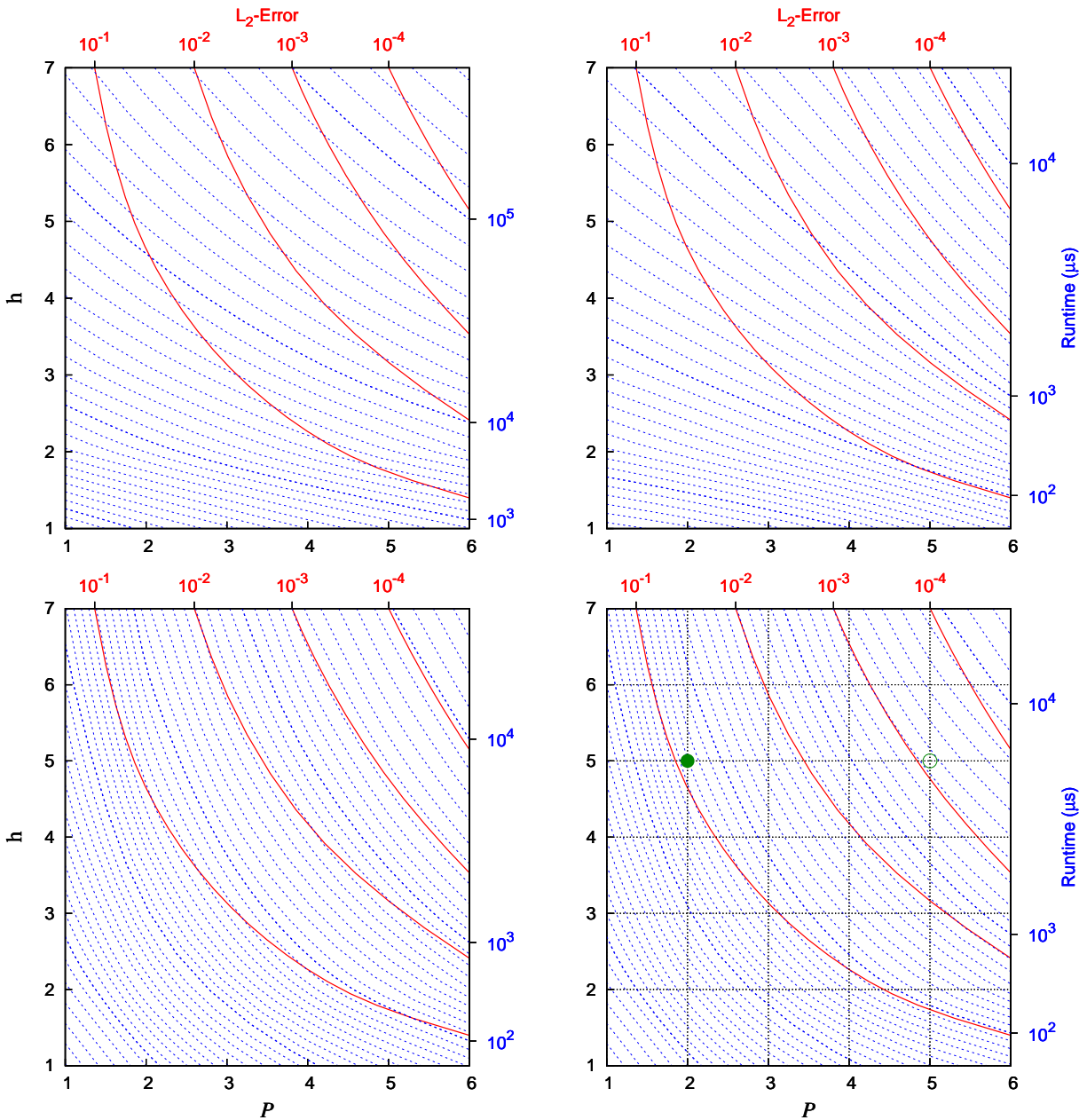


Figure 5: Contour plots showing the runtime (dotted lines) and  $L_2$ -error (solid lines and fixed across all plots) for each  $(h, P)$ -combination in solving the Helmholtz problem using tetrahedral elements. The three evaluation strategies are shown: sum-factorisation (a), elemental matrices (b) and global matrix (c). A comparison with the optimal strategy chosen for each discretisation is shown in (d), where the filled circle marks the optimal discretisation to attain a solution with a 10% error tolerance, while the open circle indicates the optimal discretisation for 0.1%.

careful selection of the discretisation influences performance. The Helmholtz problem provides an effective demonstration of the complexities and caveats of choosing discretisation and implementation to attain the maximum performance. The discrete nature of the parameters in this process also complicates matters and hinders a precise rule being formed. Nevertheless, the results do provide valuable guidelines to the approach which should be followed and significant conclusions may still be drawn from these figures.

The fourth plot in each of Figs. 4 and 5 is the culmination of this analysis where the lowest runtime across the three implementations is selected for each  $(h, P)$ -combination. We can directly read off the optimal discretisation with an error tolerance of 10%. For both hexahedral and tetrahedral elements the optimal strategy (which corresponds to a global matrix evaluation in this case) is to use linear finite elements to represent the solution. Interestingly, this is in contrast to what was found in the analysis of the two-dimensional case [15]. If instead we needed a much higher accuracy of 1%, the answer is less obvious and we should take more care. For hexahedral elements, the minimal runtime along the contour is approximately  $10^{2.2}$ . Of course, we are in fact limited to discrete choices of  $h$  and  $P$  and consequently  $h = 4, P = 3$  is the optimal choice here with a runtime of  $10^{2.4}$ . We also quote the optimal choice for 0.01% accuracy as being  $(h, P) = (5, 4)$ . Note that this has a lower runtime than  $(h, P) = (4, 6)$  even though the latter lies directly on the desired error contour.

Finally, we note some limitations of the results presented. It must be remembered that this study, as with any other in which measurements of performance are considered, is both problem and hardware dependent. The exact optimal discretisation determined through this analysis may change when the analysis is performed on a range of alternative architectures. Additionally, different problems may favour particular implementation strategies leading to a performance shift. It is of interest to quantify the extent to which different problems, for example those without infinitely smooth solutions, affect the selection of the optimal discretisation. The  $h$  and  $P$  upper-bounds in this study were governed by the memory requirements imposed by the global direct solve implementation. The use of an iterative solver would allow an extension to much higher-order expansions with a larger number of elements and allow a greater parameter space to be explored. However, the process and general trends presented here should help guide the selection process and highlight the essential need for modular codes which support both low- and high-order discretisations.

## Acknowledgements

CDC acknowledges support from the Wolfson Foundation. SJS acknowledges support from the EPSRC Advanced Research Fellowship.

## References

- [1] P.E. Bernard, J.F. Remacle, R. Comblen, V. Legat, and K. Hillewaert. High-order discontinuous Galerkin schemes on general 2D manifolds applied to the shallow water equations. *Journal of Computational Physics*, 2009.
- [2] C. D. Cantwell, S. J. Sherwin, R. M. Kirby, and P. H. J. Kelly. From h to p efficiently: strategy selection for operator evaluation on hexahedral and tetrahedral elements. *Computers and Fluids*, (submitted), 2010.
- [3] M. O. Deville, P. F. Fischer, and E. H. Mund. *Higher Order Methods for Incompressible Fluid Flow*. Cambridge University Press, Cambridge, 2002.
- [4] M. Dubiner. Spectral methods on triangles and other domains. *J. Sci. Comp.*, 6(4):345–390, 1991.
- [5] D. Gottlieb and S.A. Orszag. *Numerical analysis of spectral methods: theory and applications*. Society for Industrial Mathematics, 1977.
- [6] JS Hesthaven and T. Warburton. Nodal High-Order Methods on Unstructured Grids:: I. Time-Domain Solution of Maxwell’s Equations. *Journal of Computational Physics*, 181(1):186–221, 2002.
- [7] T. J. R. Hughes. *The Finite Element Method*. Prentice-Hall, New Jersey, 1987.
- [8] G. E. Karniadakis and S. J. Sherwin. *Spectral/hp element methods for computational fluid dynamics*. Oxford University Press, Oxford, second edition edition, 2005.
- [9] U. Lee. *Spectral element method in structural dynamics*. Wiley, 2009.
- [10] S. A. Orszag. Spectral methods for problems in complex geometries. *Advances in computer methods for partial differential equations- III*, pages 148–157, 1979.
- [11] A.T. Patera. A spectral element method for fluid dynamics: laminar flow in a channel expansion. *J. Comput. Phys.*, 54(3):468–488, 1984.
- [12] S. J. Sherwin and G. E. Karniadakis. Tetrahedral hp finite elements: Algorithms and flow simulations. *J. Comput. Phys.*, 124:14–45, 1996.
- [13] S.J. Sherwin. Hierarchical hp finite elements in hybrid domains. *Finite Elements in Analysis and Design*, 27(1):109–119, 1997.
- [14] B.F. Smith, P. Bjorstad, and W. Gropp. *Domain decomposition: parallel multilevel methods for elliptic partial differential equations*. Cambridge Univ Pr, 2004.
- [15] P. E. J. Vos, S. J. Sherwin, and M. Kirby. From h to p efficiently: Implementing finite and spectral/hp element discretisations to achieve optimal performance at low and high order approximations. *J. Comput. Phys.*, (in press), 2010.

- [16] OC Zienkiewicz, RL Taylor, and JZ Zhu. *The finite element method: its basis and fundamentals*. Elsevier Butterworth Heinemann, 2005.