V European Conference on Computational Fluid Dynamics
ECCOMAS CFD 2010
J. C. F. Pereira and A. Sequeira (Eds)
Lisbon, Portugal,14-17 June 2010

FROM H TO P EFFICIENTLY: IMPLEMENTATION OF LOW- AND HIGH- ORDER SPECTRAL/HP ELEMENT DISCRETISATIONS IN TWO- AND THREE-DIMENSIONS.

Chris D. Cantwell*, Peter E. J. Vos†, Spencer J. Sherwin† and Robert M. Kirby††

*Department of Mathematics,
Imperial College London,
London, SW7 2AZ
e-mail: c.cantwell@imperial.ac.uk
†Department of Aeronautics,
Imperial College London,
London, SW7 2AZ
e-mail:{p.vos,s.sherwin}@imperial.ac.uk
††School of Computing, University of Utah, Salt Lake City, Utah, USA.
e-mail: kirby@cs.utah.edu

Key words: spectral/hp element, high-order methods, optimisation, sum-factorisation.

Abstract. *A spectral/hp element discretisation permits both geometric flexibility and beneficial convergence properties to be attained simultaneously. The choice of elemental polynomial order ($1 \leq P \leq 15$) has a profound effect on the efficiency of different implementation strategies with their performance varying substantially for low- and high-order spectral/hp discretisations. We examine how careful selection of the strategy minimises computational cost across a wide-range of polynomial orders in two-dimensions and show how different operators, and the choice of element shape, lead to different break-even points between the implementations. We demonstrate that for low polynomial orders, a global approach is typically the most efficient, while for high-order polynomials a sum-factorisation technique, exploiting the tensor-product structure of elemental expansions, is optimal. For intermediate orders, one may find that using local-elemental matrices are more desirable.*

In three-dimensions, higher expansion orders quickly lead to a large increase in the number of element-interior modes, particularly in hexehedral elements. For a typical boundary-interior modal decomposition, this can rapidly lead to a poor performance from a global approach. We determine performance boundaries for the different implementations in 3D geometries. Furthermore, increased memory requirements may cause an implementation to show poor run-time performance, even if the strict operation count is minimal, due to detrimental caching effects and other machine-dependent factors.

1 INTRODUCTION

Spectral/hp element solvers are now a popular choice in many areas such as fluid dynamics and structural mechanics. These research communities benefit from both the flexibility offered by the elemental decomposition of the domain and the high accuracy and preferential convergence properties of a spectral method. Unlike conventional finite-element or pure spectral techniques spectral/hp methods permit greater scope for optimisation through mesh element density, polynomial order and evaluation strategies of the various numerical operators. The importance of choosing the correct strategy should not be underestimated, as is evident from the results presented in this paper. Even in two dimensions, changing the polynomial order by just one order may incur a runtime performance penalty on the order of 50% if the strategy is not also switched to the optimal choice for the new order.

In formulating a spectral/hp element method a geometrically complex domain is divided into a tesselation of $\mathcal{K}$ non-overlapping elements on which the solution field is expressed in terms of polynomial expansions up to a chosen order, P. The typical choice for P varies between communities, with those from the finite-element community using expansions up to fourth-order[2], and those in the spectral/hp element community typically considering polynomials up to 15th order[3, 7]. For the purposes of this study we consider polynomial orders in the broad range of $1 \leq P \leq 15$.

The elemental regions may be of different shapes, each with their own associated performance characteristics. For two-dimensional problems, elements are typically triangular or quadrilateral, while in three dimensions elements may be hexahedral, tetrahedral, prismatic or pyramidal. This gives great geometric flexibility but may lead to complexities in choosing an optimal evaluation strategy due to the differing performance of expansions on each of these elemental regions. In two or more dimensions, the elemental modes may be chosen to form a modal basis constructed as a tensor product of one-dimensional polynomials. This permits the use of strategies which dramatically reduce the operation count, and therefore improve the overall performance, of a given operator when compared with a naive local matrix implementation.

These elemental expansions are then extended to a global context, typically with the enforcement of a continuity constraint. An invertible assembly matrix describes the scattering of global modes onto the local elemental coefficients allowing the problem, and therefore the evaluation of the operators, to be approached in either a global or elemental context. Elemental evaluation in two- or three-dimensions may then be performed using a local matrix, or by exploiting the tensorial nature of the expansions and using a sum-factorisation approach[5]. The formulation of each of these methods is detailed in the next section.

2 SPECTRAL/HP FORMULATION AND EVALUATION STRATEGIES

In this section we briefly describe the construction of the spectral/hp element framework and outline the various evaluation strategies available for computing four commonly-used operators. The operators we consider are the backward transform, inner product, mass matrix and Helmholtz operators. Expansion bases may be either nodal or modal in nature, but we will

focus on the modal case throughout due to the additional scope for optimisation available from a tensorial construction. We first define the various elemental shapes in reference space which are subsequently mapped onto the physical elements which constitute the domain.

2.1 Two- and three-dimensional expansion bases

We define the standard quadrilateral element as the square $\mathcal{Q}^2 = \{(\xi_1, \xi_2) \in [-1, 1] \times [-1, 1]\}$ on which the physical solution u may be represented as an expansion in terms of two-dimensional modes ϕ_n as

$$u(\xi_1, \xi_2) = \sum_{n \in \mathcal{N}} \phi_n(\xi_1, \xi_2)\hat{u}_n = \sum_{p=0}^{P}\sum_{q=0}^{P} \psi_p(\xi_1)\psi_q(\xi_2)\hat{u}_{pq} \tag{1}$$

Here the $\{\phi_n\}$ are the tensor product of the one-dimensional expansions $\{\psi_p\}$ and $\{\psi_q\}$.

The triangular element $\mathcal{T}^2 = \{(\xi_1, \xi_2) : -1 \leq \xi_1, \xi_2; \xi_1 + \xi_2 = 0\}$ does not immediately expose a straightforward tensorial construction. However, the region may be defined in terms of a collapsed coordinate system[1],

$$\eta_1(\xi_1, \xi_2) = 2\frac{1 + \xi_1}{1 - \xi_2} - 1,$$
$$\eta_2(\xi_1, \xi_2) = \xi_2,$$

as $\mathcal{T}^2 = \{(\eta_1, \eta_2) \in [-1, 1] \times [-1, 1]\}$. This restores the tensorial nature of the expansion, but introduces a dependency of the η_2 modes on the η_1 modes. The solution on a triangular element[6] is therefore not a full tensorial expansion and is expressed as

$$u(\xi_1, \xi_2) = \sum_{n \in \mathcal{N}} \phi_n(\xi_1, \xi_2)\hat{u}_n = \sum_{p=0}^{P}\sum_{q=0}^{f(p)} \psi_p(\eta_1(\xi_1, \xi_2))\psi_{pq}(\eta_2(\xi_1, \xi_2))\hat{u}_{pq} \tag{2}$$

The standard hexehedral element is $\mathcal{Q}^3 = \{(\xi_1, \xi_2, \xi_3) \in [-1, 1]^3\}$. The modal expansion is therefore a straightforward extension of the quadrilateral case.

$$u(\xi_1, \xi_2, \xi_3) = \sum_{n \in \mathcal{N}} \phi_n(\xi_1, \xi_2, \xi_3)\hat{u}_n = \sum_{p=0}^{P}\sum_{q=0}^{P}\sum_{r=0}^{P} \psi_p(\xi_1)\psi_q(\xi_2)\psi_r(\xi_3)\hat{u}_{pqr} \tag{3}$$

The expansion basis on each local element region is defined by a mapping from the basis of the standard element of corresponding shape via a mapping $\chi_i^e : \Omega_{st} \rightarrow \Omega_e$ defined as

$$x_1 = \chi_1^e(\xi_1, \xi_2) \qquad x_2 = \chi_2^e(\xi_1, \xi_2).$$

2.2 Global assembly

A global formulation requires a degree of connectivity between the individual elements, typically in the form of a C^0 continuity condition. The boundary/interior decomposition of the elemental modes simplifies this aspect since all interior elemental modes are themselves global modes. It is then only necessary to be concerned with the elemental boundary modes when constructing a global formulation.

Given our tesselation of $\mathcal{K}$ elements, each with $\mathcal{N}$ elemental modes, the solution on the entire domain can be represented as

$$\begin{aligned} u(x_1, x_2) &= \sum_{m \in \mathcal{N}^g} \Phi_m(x_1, x_2) \hat{u}^g_m \\ &= \sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{N}} \phi^k_n(x_1, x_2) \hat{u}^k_n. \end{aligned}$$

The transition to global degrees of freedom can be succinctly expressed as a highly sparse local-to-global mapping matrix, $\mathcal{A}$, mapping the $\mathcal{K} \times \mathcal{N}$ elemental degrees of freedom onto the $\mathcal{N}^g$ global degrees of freedom,

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

2.3 Evaluation strategies

With the foundations of our spectral/hp formulation in place we may now consider the different approaches to evaluating numerical operators in our scheme. We extend the two-dimensional results[8] to three dimensions and consider the following three strategies

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

In the global context a sparse $\mathcal{N}^g \times \mathcal{N}^g$ matrix is constructed which directly solves for the global coefficients. This approach is typical of finite element formulations in which all modes are essentially elemental boundary modes and results in a significantly lower operation count than handling each element individually at lower polynomial orders.

The remaining two evaluation strategies are performed at the elemental level. The global modes $\hat{\boldsymbol{u}}^g$ are scattered onto the local modes $\hat{\boldsymbol{u}}^l$ upon which an elemental evaluation

$$\hat{y}^e_m = \sum_{n \in \mathcal{N}} a_e(\phi^e_m, \phi^e_m) \hat{u}^e_n \quad \forall (m, e) \in (\mathcal{N}, \mathcal{K})$$

is performed. Here we use $a_e(v, u)$ to represent a general bi-linear operator typical of when a PDE is expressed in weak form through a Galerkin construction. The resulting vector, $\hat{\boldsymbol{y}}$ is then reassembled to give the global solution.

The sum-factorisation strategy[5] exploits the tensorial nature of the elemental modes and the performance benefits from this construction have been appreciated throughout the history of spectral elements[4]. Details of the formulation for quadrilateral and triangular elements are given in[8], and so we restrict ourselves to presenting the formulation in the hexahedral region. For simplicity we will consider just the backward transform in detail. The other operators may be expressed in a similar form. The expansion basis defined in Equation (3) can be reorganised as follows

$$u(\xi_{1i},\xi_{2j},\xi_{3k}) = \sum_{p=0}^{P}\psi_p(\xi_{1i})\left\{\sum_{q=0}^{P}\psi_q(\xi_{2j})\left\{\sum_{r=0}^{P}\hat{u}_{pqr}\psi_r(\xi_{3k})\right\}\right\}. \tag{4}$$

Further to the reduction in operation count, the summation can be expressed as a sequence of matrix-matrix multiplications. To see this we define the order of the elemental modes $\hat{u}_{pqr}$ in the vector $\hat{\boldsymbol{u}}$ to be such that p (rows) runs fastest, followed by q (columns), and r (stacks) runs slowest. Let $\hat{\boldsymbol{U}}_{[P]}$ define a $P \times P^2$ matrix where each column corresponding to a *row* of coefficients in $\hat{\boldsymbol{u}}$; that is $\hat{\boldsymbol{u}}[iP : (i+1)P - 1]$, $0 < i < P^2$. The backward-transform operator from Equation (4) can now be expressed as a sequence of matrix-matrix multiplications,

$$\boldsymbol{U}_{[P]} = \left[\left(\hat{\boldsymbol{U}}_{[P]}^{\top}\boldsymbol{B}_{\boldsymbol{0}}^{\top}\right)^{\top}\boldsymbol{B}_1^{\top}\right]^{\top}\boldsymbol{B}_2^{\top} = \boldsymbol{B}_1(\hat{\boldsymbol{U}}_{[P]}^{\top}\boldsymbol{B}_0^{\top})\boldsymbol{B}_2^{\top}$$

The other operators may be expressed in a similar way. For example, the inner product operator and mass-matrix operators become

$$\hat{\boldsymbol{U}}_{[\boldsymbol{P}]} = \left[\boldsymbol{B}_0^{\top}w(\boldsymbol{U}_{[Q]})\boldsymbol{B}_1\right]^{\top}\boldsymbol{B}_2,$$

and

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

respectively, where $w(\boldsymbol{U}_{[P]})$ applies the quadrature metrics to the coefficients.

3 OPTIMAL STRATEGY IN 2D

We first present a summary of the optimal strategy in two dimensions based on computational runtime. While a comparison of strict operation count may give an estimation of trends relating to the comparison of the different strategies, we have found in practice that a direct study of timings is necessary to identify the break-even points. With this in mind, it is also notable that the specifications of the hardware also have a significant effect on the results. One strategy deemed optimal for a given polynomial order on one particular hardware configuration may not be optimal on a different configuration. Processor cache size is particularly relevant here since performance will suffer when a given matrix operation no longer fits into cache. This is particularly evident by comparing the results presented here with those given in[8] which

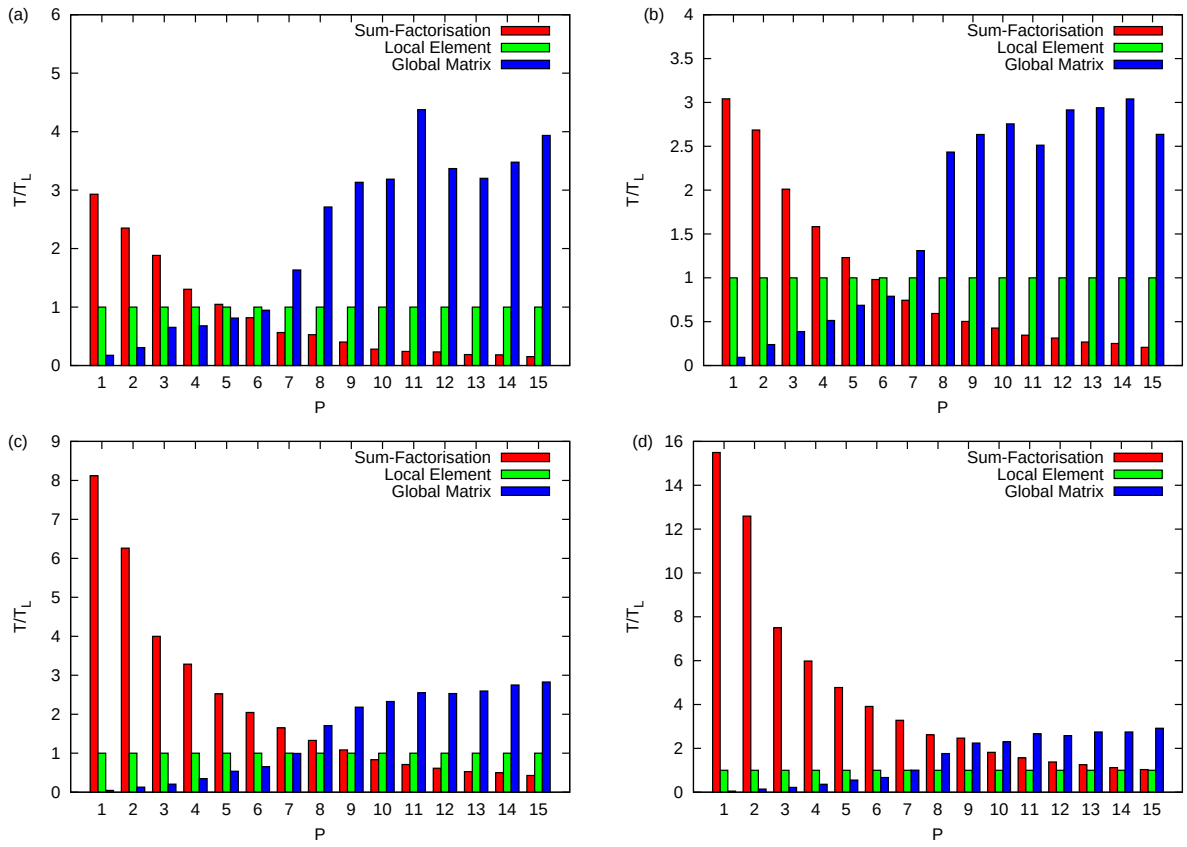

Figure 1: Comparitive performance of (a) backward transform, (b) inner product, (c) mass matrix and (d) Helmholtz operators on a mesh of 100 quadrilateral elements.

were computed on a MacBook Pro (2.33Ghz dual core processor, 4MB cache, 2GB RAM) using reference implementations of BLAS and LAPACK. For the present results all simulations were conducted on a Mac Pro (two 2.8Ghz quad-core processors, 24MB cache total, 10GB RAM) with the Accelerate Framework enabled, allowing parallelisation of certain BLAS and LAPACK operations. The substantial differences highlight the variability of performance with architecture and use of optimised libraries.

Figure 1 shows a comparison of the performance of the different strategies for quadrilateral elements. As one would expect global matrix evaluation provides the highest performance at low orders while the sum-factorisation is superior at high orders. The best approach to choose in the intermediate range is less obvious. The global strategy is consistently the optimal choice up to fifth order for all operators. However, the sum-factorisation approach suffers when evaluating the more complex mass matrix and Helmholtz operators. For these there exists an intermediate range of polynomial orders for which local elemental evaluation is the optimal strategy. The polynomial range for each operator where each strategy is optimal is summarised in Table 1.

For triangular elements, the global strategy is optimal for a larger range of polynomials than for the quadrilateral case (see Figure 2). This can be attributed to the higher ratio of

P	Global Matrix	Local Matrix	Sum-Factorisation
Bwd. Trans.	1-5	-	6-15
Inner Prod.	1-6	-	7-15
Mass Matrix	1-7	8-9	10-15
Helmholtz	1-6	7-15	-

Table 1: Table of optimal strategies selection for different operators on quadrilateral meshes.

P	Global Matrix	Local Matrix	Sum-Factorisation
Bwd. Trans.	1-6	7-8	9-15
Inner Prod.	1-7	-	8-15
Mass Matrix	1-8	9-15	-
Helmholtz	1-8	9-15	-

Table 2: Table of optimal strategies selection for different operators on triangular meshes.

boundary modes to interior modes leading to a proportionally lower operation count. Another striking feature is the poor performance of sum-factorisation in general, particularly for complex operators. This is due to the triangular expansion not being a full tensorial basis and therefore we are unable to formulate the operation purely using `dgemm` operations. This leads to a local elemental evaluation strategy being optimal at high polynomial orders. Again we summarise the break-even points in Table 2.

4 OPTIMAL STRATEGY IN 3D

We now discuss the relative performance of the different strategies in three dimensions in the context of hexahedral elements. The pure tensorial nature of these expansions would suggest that a strategy map similar to the quadrilateral case should be observed, although the relative performances would diverge more rapidly due to the number of modes increasing as $O(P^3)$.

The timing results for the hexahedral element are shown in Figure 3. Missing entries for the global strategy in this figure are due to memory constraints preventing evaluation with the global matrix. It is immediately apparent that the global matrix method is now only optimal at low-orders, when using polynomial bases up to 3rd order. This is somewhat unsurprising given the cubic growth of modes in three dimensions. At higher orders it becomes rapidly sub-optimal and the size of the global matrix makes this approach infeasible.

The sum-factorisation approach proves even more effective in three dimensions than was evident in two dimensions and is the optimal strategy, even for complex operators, above 8th order. As with quadrilateral meshes, the local matrix approach is typically only optimal for more complex operators (especially those involving derivatives), although to a lesser extent. The apparent improvement in performance of the global strategy at higher orders is in fact due to a reduction in the performance of the local matrix approach, against which the plot is

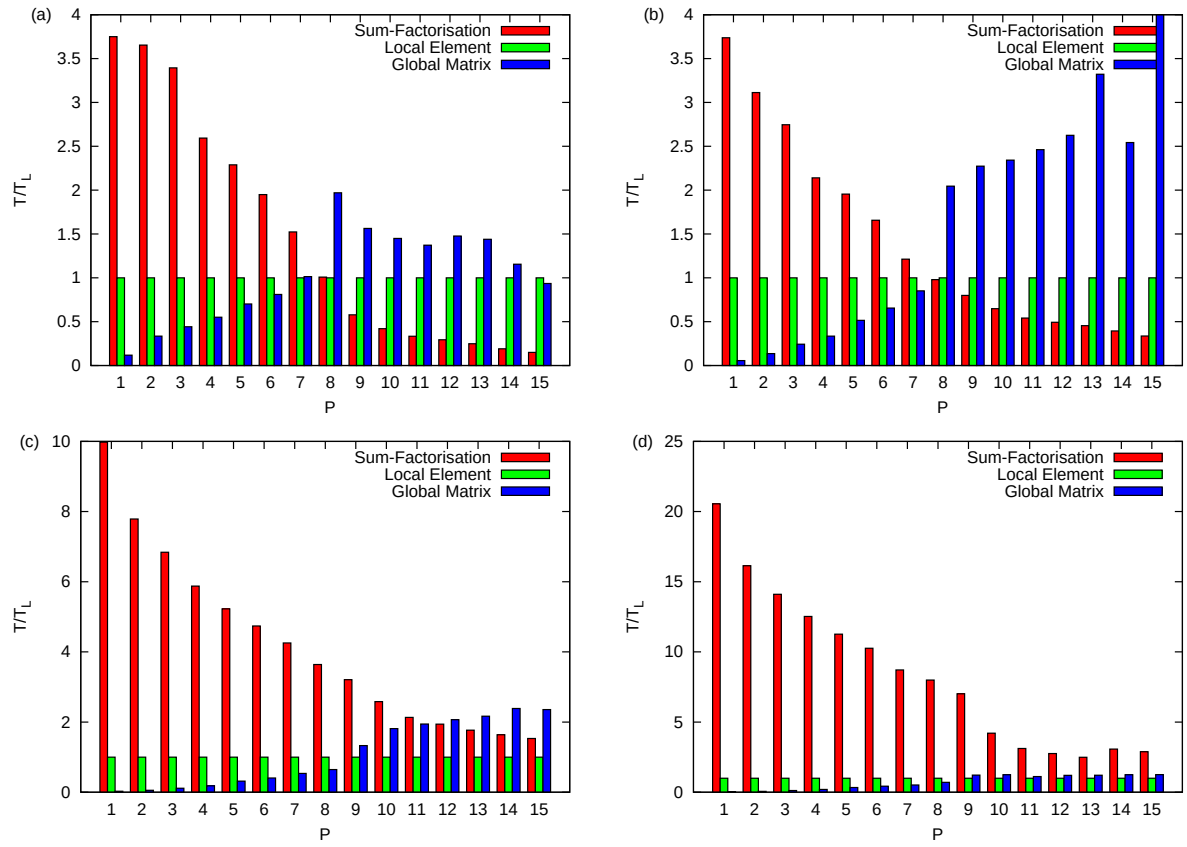

Figure 2: Comparitive performance of (a) backward transform, (b) inner product, (c) mass matrix and (d) Helmholtz operators on a mesh of 186 triangular elements.

normalised. This is likely due to machine architecture characteristics resulting in cache-misses.

5 CONCLUSIONS

The results presented here highlight the importance of choosing the optimal evaluation strategy in achieving the greatest possible performance from spectral/hp element codes. This is seen to be even more critical in moving from two-dimensional to three-dimensional problems where the number of modes increases as $O(P^3)$ and the separation between the performance of global and elemental strategies becomes more distinct.

The results align with the general philosophy, already practiced by the spectral/hp community, of using global strategies for low-order polynomial bases and exploiting the elemental and/or tensorial nature of the multi-dimensional modal expansions for high-order problems. However, it is clear that there is no distinct boundary between these choices, particularly in intermediate regimes, and the optimal strategy is dependent on both the element type and numerical operator being evaluated. The architecture of the system is also identified as playing a significant role in the performance of some strategies and therefore highlights the importance

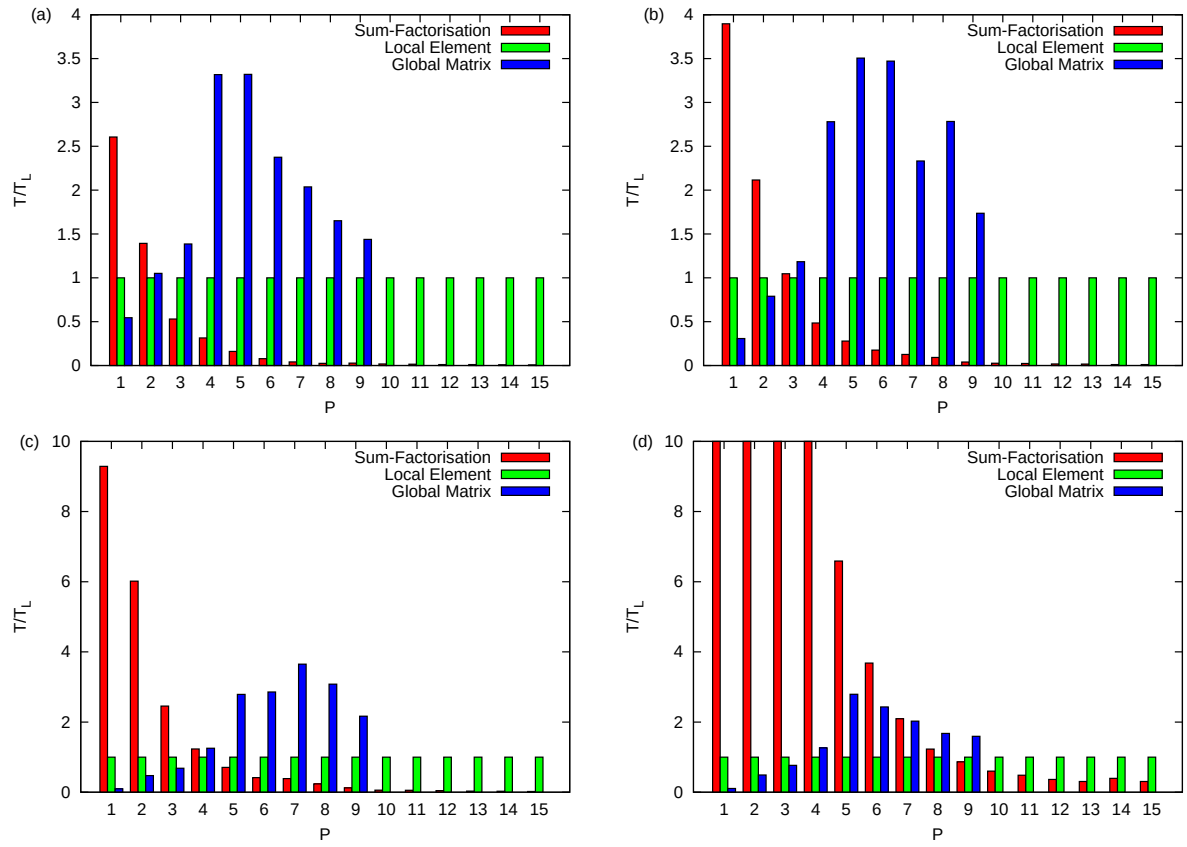

Figure 3: Comparitive performance of (a) backward transform, (b) inner product, (c) mass matrix and (d) Helmholtz operators on a mesh of 64 hexahedral elements. The runtime of the Helmholtz operator at orders 1-4 is in excess of 10 times that of the local matrix evaluation.

of tuning any particular optimisation strategy to the hardware being used. It is therefore clear that the ability of a spectral/hp element code to support a range of operator implementations will allow users to move between low- and high-order schemes efficiently on a variety of architectures.

REFERENCES

[1] M. Dubiner. Spectral methods on triangles and other domains. *J. Sci. Comp.*, 6(4):345–390, 1991.

[2] T. J. R. Hughes. *The Finite Element Method*. Prentice-Hall, New Jersey, 1987.

[3] G. E. Karniadakis and S. J. Sherwin. *Spectral/hp element methods for computational fluid dynamics*. Oxford University Press, Oxford, second edition edition, 2005.

[4] J. M. Melenk, K. Gerdes, and C. Schwab. Fully discrete hp-finite elements: fast quadrature. *Comp. Meth. Appl. Mech. & Engng*, 190(32-33):4339–4364, 2001.

P	Global Matrix	Local Matrix	Sum-Factorisation
Bwd. Trans.	1	2	3-15
Inner Prod.	1-2	3	4-15
Mass Matrix	1-3	4	5-15
Helmholtz	1-3	4-8	9-15

Table 3: Table of optimal strategies selection for different operators on hexahedral meshes.

[5] S. A. Orszag. Spectral methods for problems in complex geometries. *Advances in computer methods for partial differential equations- III*, pages 148–157, 1979.

[6] S. J. Sherwin and G. E. Karniadakis. A triangular spectral element method; applications to the incompressible navier-stokes equations. *Comp. Meth. Appl. Mech. & Engng*, 123(1-4):189–229, 1995.

[7] B. Szabó and I Babuška. *Finite Element Analysis*. John Wiley & Sons, New York, 1991.

[8] 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.