# Characterizing the Accuracy of Summation-by-Parts Operators for Second-Derivatives with Variable-Coefficients

by

Guang Wei Yu

Supervisor: D. W. Zingg

April 13, 2013

## Abstract

This paper presents the characteristics of solution accuracy as mesh is refined for the second-order summation-by-parts(SBP) operator with simultaneous approximation terms (SATs) for second-derivatives with variable-coefficients through numerical experiments. Second-order parabolic partial differential equations (PDEs) with constant and variable-coefficients on the second derivative are studied. An approximately 70% reduction of error is demonstrated in SBP-SAT operators with optimized parameters over SBP-SAT operators without optimization in parameter for the specific types of problems. Convergence trends are outlined for different solution shapes with constant and variable-coefficients, and similarly for different variable-coefficients with matching solution shapes. The types of functions examined include sinusoidal, exponential, hyperbolic cosines, and Gaussian functions. We use parameter p defined by the fact that theoretical interior order of accuracy is 2p to notate convergence rate. Furthermore, we distinguish the operators as compact-stencil if it uses minimum stencil width to approximate the continuous form of the discrete operator, and wide-stencil otherwise. We observed that the convergence rate can be lower (global order p+1 instead of p+2 in theory) for the compact-stencil operators depending on variable-coefficient and solution shapes. We also observed that convergence rate can be higher than theoretical values for wide-stencil operators (p+2 instead of p+1 in theory) for specific cases.

## Acknowledgments

I would like to express my deepest gratitude to my thesis advisor Dr. David W. Zingg for his guidance, care, and patience. I would also like to thank David Fernández PhD(ABD) who patiently helped me throughout the project. Finally, I would like to thank my parents who have provided moral support and always encouraged me with their best wishes.

# Contents

# List of Figures

# 1 Introduction

The study of computational fluid dynamics (CFD) is motivated by physical phenomena occurring in flow of fluids near an object, such as shock waves, slip surfaces, boundary layers, and turbulence[3]. The underlying governing equations are the Navier-Stokes equations, which appear as second-order partial differential equations (PDEs) with variable-coefficients, explaining and relating the interactions of these phenomena. Many of these relations have no analytic solution due to their non-linear nature, and motivate the use of numerical flow solvers[3]. In application, solving these equations is of interest to engineers dealing with design and optimization problems such as aerodynamic shape optimization [9], and of interest to scientists for research and understanding purpose. For a numerical method to be useful for these applications, its efficiency is of utmost importance, and is measured by the accuracy achieved relative to the cost of obtaining the solution [3]. Therefore paramount emphasis has been placed on improving efficiency, making this an active area of research.

For most finite difference solvers, which are dominantly used in application involving well-structured meshes such as airfoil shape optimization, it is more computationally efficient to adopt higher-order operators with coarser mesh than low-order operators with fine mesh [5]. Furthermore, multi-block adaptation of the finite difference is desirable for many complex problems [6]. A discretization that best complements the solver when adopting both these methods is the summation-by-parts (SBP) operators with simultaneous approximation terms (SAT) [4].

First derivative SBP-SAT operators are well developed and characterized [11] but not much work has been done on the second derivative operators complying to the SBP-SAT formulation, until Mattsson developed a successful second derivative SBP-SAT operator [8]. This is later extended to account for variable-coefficient [7, 12] which is especially of interest due to its usefulness for solving Navier-Stokes equations. This has pushed similar operators to be developed for more specified application such as multi-disciplinary optimization (MDO) of aircraft by Fernández and Zingg [2]. However, characterization of the efficiency of these second derivative SBP-SAT operators applied to specific problems does not exist, and developing such framework is an important task that remains to be done before the operators can be meaningfully used.

This thesis focuses on applying SBP-SAT operators to solve second-order PDEs with variable-coefficient and characterizing the accuracy obtained for a variety of solutions. In particular, this thesis emphasize on the formulation presented by Fernández and Zingg [2]. The

aim is to examine, through numerical simulation, the efficiency of the given SBP-SAT methods for different cases of solutions and variable-coefficient choices for a second-order PDE, and in turn to obtain understanding in choosing best operator parameters for these specific problems.

## 2 Literature Review

There are two main motivations behind adopting SBP-SAT discretization for finite difference methods, which are its compatibility with multiblock finite difference approach and the its ability to generate higher-order discretization that properly address boundary treatment [4, 5, 6].

For efficient flow solvers, multiblock finite difference approach is often adopted as they break up computational domain into sub-domains to allow parallel solution algorithm [6]. However, this approach poses challenges at block interfaces such as singularities arising from geometry of the block and interdependencies in mesh refinement of the blocks [4]. One approach to this problem is through the use of summation-by-parts (SBP) operators which are finite-difference operators that mimic integration by parts, combined with simultaneous-approximation-terms (SAT) [4]. The SBP-SAT formulation can be used to eliminate the challenges by requiring only $C^0$ continuity at block interface and properly treating exceptional point while having time-stability and low communication overhead compared to halo approaches. The motivation for SAT terms was originally to enforce time-stable boundary conditions [1], but the method has been developed to apply to block interfaces [10], enforcing weak boundary conditions and coupling blocks by penalty terms [6].

In addition to adopting multi-block approach, and more importantly in our context, Kreiss and Oliger showed the use of higher-order finite-difference approximations were much more computationally efficient [5], at the cost of a more mathematically complex method that exhibit difficulty in boundary treatment. The computational efficiency comes from the fact that despite having more computational cost per grid node, higher-order methods reduce the total computational cost by enabling the use of coarser meshes to achieve given level of numerical accuracy [2]. Again, to address the difficulty in boundary treatment, SBP-SAT formulation can be used to systematically generate higher-order operators that retain the advantages of second-order schemes [4] such as $C^0$ continuity requirement at mesh interfaces, time stability, and exceptional point treatment without increasing communication overhead at block interfaces.

Mattsson developed the first practical formulation for second-order PDEs [8] and the result is extended to account for variable-coefficient by Mattsson [7] and Zingg [12] which is useful in

solving Navier-Stokes equations. In his recent work, Mattsson presents compact-stencil SBP operators up to $5^{\text{th}}$ order accuracy. [7]. Furthermore, recent progress in developing similar methods by Fernández [2] focus on deriving compact-stencil operators that generalizes to higher order and exhibit more appropriate accuracy characteristics for specific applications, and validation is only done using simple linear convection-diffusion equation.

Currently there lacks a more in-depth examination of operator characteristics subject to specific choices of variable-coefficient as well as solution shape. Additionally, not much is known about how the operators can be tuned to become more efficient in these specific applications. There is a present interest in characterizing the SBP-SAT discretized operators for various types of problems to establish a basis for performance which can be used to compare with related operators, as well as optimizing the operator. The main objective of this thesis is to characterize the SBP-SAT operators subject to various variable-coefficient choices on the second derivatives, with respect to accuracy, which will help address the former interest. The thesis will also compare optimized operator in performance to establish if the SBP-SAT operators can be optimized.

# 3 Theory

The problem to which we apply the SBP-SAT discretization is a second-order PDE of the following form

$$\frac{\partial u}{\partial t} = -a \left( \frac{\partial u}{\partial x} \right) + \varepsilon \frac{\partial}{\partial x} \left( b(x) \frac{\partial u}{\partial x} \right), a > 0, \varepsilon > 0 \tag{1}$$

with proper boundary and initial conditions.

In order to begin the analysis process, we outline the components of SBP-SAT method. The SBP-SAT method can be broken into the SBP operator and the SATs, with the SBP operator ignores boundary and acts on the interior nodes only and SATs weakly enforcing boundary conditions.

For first derivative SBP operator $D_1$, $D_1$ must approximate first derivative, and additionally be of the form $D_1 = H^{-1}\Theta$, for a positive-definite diagonal matrix $H \in \mathbb{R}^{(N+1)\times(N+1)}$ and $\Theta$ satisfying the SBP property $\Theta + \Theta^T = diag(-1, 0, \ldots, 0, 1)$.

For second derivative SBP operator, we start with constant-coefficient case and extend to variable-coefficient operators. The second derivative SBP operator for the constant-coefficient

case, $D_2 \in \mathbb{R}^{(N+1)\times(N+1)}$, is defined as $D_2(B) = H^{-1}\{-M + EBD_b\}$, where $E = diag(-1,0,\ldots,0,1), B = diag(\beta_0,\ldots,\beta_N), D_b$ is approximation to the first derivative at boundaries, and $M = D_1^T HBD_1 + R$. $M$, $R$, and $B$ are positive-semi-definite, while $M$ and $R$ are also symmetric. $H$ used here is same as the $H$ in first derivative, and B is discretization of the variable-coefficient $b(x)$ that appear in equation 1.

These constraints form a system of nonlinear equations that can be solved for many different SBP operators including those currently being worked on by Fernández, which is what this paper will focus on characterizing.

The left and right hand side SAT terms for discretization of the boundary treatment defined by Fernández is of the form

$$SAT_L = \tau_L H^{-1}\mathbf{e_L}(\alpha_0 q_0 + \beta_0 b_0 (BD_b\mathbf{u})_0 - g_0)$$
$$SAT_R = \tau_R H^{-1}\mathbf{e_R}(\alpha_1 q_1 + \beta_1 b_1 (BD_b\mathbf{u})_1 - g_1)$$

(2)

with $\mathbf{u}$ being discretization of solution $u(x)$ to the PDE problem stated in equation 1 ($u$) and subscript 0 and 1 being notation for evaluation at left and right boundary respectively. We also have $e_L = [1,0,\ldots,0,0]^T$ and $e_R = [0,0,\ldots,0,1]^T$ for $e_L, e_R \in \mathbb{R}^{N+1}$. Fernández further requires $\tau_L = \frac{\varepsilon}{\beta_0}$ and $\tau_R = -\frac{\varepsilon}{\beta_1}$, as well as the following inequalities for the problem to be well-posed

$$a + \frac{2\varepsilon\alpha_0}{\beta_0} < 0$$
$$a + \frac{2\varepsilon\alpha_1}{\beta_1} > 0$$

(3)

The semi-discrete form of equation 1 is

$$\frac{d\mathbf{u}}{dt} = -aD_1\mathbf{u} + \varepsilon H^{-1}\{-M + BED_b\}\mathbf{u} + SAT_L + SAT_R$$

(4)

and in this paper we further restrict ourselves to concern with only the steady state solution when $\frac{\partial u}{\partial t} = 0$ or $\frac{d\mathbf{u}}{dt} = 0$. We will call this the model PDE which we will pick the variable-coefficient $b(x)$ and perform discretization appropriately. With regards to operators by Fernández, we also choose $\beta_0 = -1$ and $\beta_1 = 1$ at boundary, and characteristic boundary conditions at which $\alpha_0 = 1$ and $\alpha_1 = 0$. Furthermore, we pick following values for the model PDE coefficients: $a = 1$ and $\varepsilon = 1$. This gives us the following model PDE (5) with the SATs defined exactly as equation 2.

4

$$0 = -D_1\mathbf{u} + H^{-1}\left\{-M + BED_b\right\}\mathbf{u} + SAT_L + SAT_R$$
$$SAT_L = \tau_L H^{-1}\mathbf{e_L}(q_0 - 1b_0(BD_b\mathbf{u})_0 - g_0) \tag{5}$$
$$SAT_R = \tau_R H^{-1}\mathbf{e_R}(b_1(BD_b\mathbf{u})_1 - g_1)$$

For the constant-coefficient case, we will analytically solve the PDE and compare error between numerical and analytical solution. For the variable-coefficient cases, we will use the method of manufactured solution and assume some solution u(x) with properties that we are interested in studying. We introduce appropriate source terms to satisfy our discretized model PDE.

Finally, we make a note that in this paper, we will adopt the following notation when talking about operator order of accuracy: we will refer to the operator as $O(i,j,k)$ where $i$ is the interior order of accuracy, $j$ is the boundary order of accuracy, and $k$ is the global order of accuracy.

# 4   Methods

We start with investigating convection equation which involves the continuous PDE with periodic boundary condition and initial value specified by equation 6 with spatial discretization using periodic operator.

$$\left(\frac{\partial u}{\partial t}\right) + \left(\frac{\partial u}{\partial x}\right), 0 \leq x \leq 2\pi, u(0,t) = u(2\pi,t), u(x,0) = u_0(x) \tag{6}$$

In particular, we examined the discretization using second-order operator giving us semi-discretized equation 7,
$$\frac{\partial u}{\partial t} = -\frac{1}{2\Delta x}B_p(-1,0,1)\mathbf{u} \tag{7}$$
where $B_p(-1,0,1)$ is the second-order periodic operator, which when expressed in matrix form, is composed of second-order centered difference scheme along the diagonal with -1 at $(1,n)$ and 1 at $(n,1)$ entries of the matrix, and $\Delta x$ is the mesh spacing. We initially used 4th order Runge-Kutta time marching scheme to obtain stable solution. However this method incurred some difficulties as we moved to second-order operator and we employed the method of solving linear systems of equation via matrix inversion, which is explained in following paragraphs.

Moving onto the model PDE involving second-order derivatives, which is the focus of this thesis, we deal with the model PDE presented in equation 5. We use existing operator schemes

to generate the components $\{D_1, H, M, B, E, D_b\}$ in equation 5. Dealing with only problems that involve scalar variable values at mesh nodes, we can express the solution $\mathbf{u}$ as $n \times 1$ vector where $n$ is the number of mesh nodes. The operators are then $n \times n$ matrices, and we will represent second derivative operator as $D_2 = H^{-1}\{-M + BED_b\}$. Further more, we will move away from using Runge-Kutta scheme and start with assumed steady state. We will build the operators fully in matrix form, and proceed to invert them to give us the solution $\mathbf{u}$ directly. We will employ a source term to allow us to assign arbitrary variable-coefficient $B$ and solution $\mathbf{u}$. This source term will allow the continuous form of the model PDE to be satisfied (right hand side of equation 1 will go to 0 when added to the source term, when $a = 1, \varepsilon = 1$). This implies source term will take on the continuous form of 8, and discretized form $SOURCE$ is simply the vector of values evaluated at each mesh nodes.

$$source = -\left(\frac{\partial u}{\partial x}\right) + \frac{\partial}{\partial x}\left(b(x)\frac{\partial u}{\partial x}\right) \tag{8}$$

This gives us linear system of equation 9

$$0 = -D_1\mathbf{u} + D_2\mathbf{u} + SAT_L + SAT_R + SOURCE \tag{9}$$

, which we can invert and solve for the solution $\mathbf{u}$ as equation 10

$$\mathbf{u} = (D_1 - D_2)^{-1}(SAT_L + SAT_R + SOURCE) \tag{10}$$

We take the convergence rate to be the error between analytic solution and numerical solution given by equation 10. The error is defined using the following norms:

$$\|\mathbf{e}\|_2 = \sqrt{\frac{\sum\limits_{i=1}^{n} \left|u_{numerical,i} - u_{analytic,i}\right|^2}{n}} \tag{11}$$

$$\|\mathbf{e}\|_\infty = \max_{1 \le i \le n}\left|u_{numerical,i} - u_{analytic,i}\right| \tag{12}$$

where $n$ is the number of mesh nodes and $\mathbf{u_{analytic}}$ is the analytical solution evaluated at the $n$ nodes. The terms $u_{analytic,i}$ and $u_{numerical,i}$ are the value of $\mathbf{u_{analytic}}$ and $\mathbf{u_{numerical}}$ evaluated at $i$th node respectively.

Numerical experiments were conducted by varying solutions and coefficients. The error defined

using norms in (11) and (12) is plotted against mesh size both in log domain. The slope is computed using linear interpolation of a range of points specified manually upon examining the convergence plot. The slope is compared to theoretical global order of accuracy. The error magnitudes are examined approximately and average difference is noted for significant results as in the case of optimized operator.

Only fourth order results are examined for the different compact-stencil operators due to limitation in operators generated.

# 5 Results and Discussion

Results are divided into the constant-coefficient case (section 5.1) and the variable-coefficient cases (section 5.2). In the variable-coefficient cases, the effect of varying solution choice, varying the variable-coefficient, and performance of operators are discussed respectively. The convergence trend values for selected cases are included in appendix A.

## 5.1 Constant-Coefficient

We obtained convergence rate of solution error described by $\|\mathbf{e}\|_2$ norm and $\|\mathbf{e}\|_\infty$ norm to compare between the wide-stencil operator and compact-stencil operator by Fernández for the constant-coefficient case. The results for second-order interior accuracy SBP-SAT operators are presented in Figure 1. The prescribed solution is the analytical solution for PDE which is $u(x) = \frac{e^{px}-1}{e^p-1}$, this implies that no source terms are required.

The convergence plots have sections emphasized to indicate the portion of data used to interpolate for convergence rate. $\|\mathbf{e}\|_2$ is the $\|\mathbf{e}\|_2$ norm of the error in solution, and $\|\mathbf{e}\|_\infty$ is the infinite norm of the error in solution using the SBP-SAT operators and comparing to the exact solution.

The result in Figure 1 agrees with the convergence rate proposed by Fernández [2]. More importantly, we also show that extending to higher order constant-coefficient operators ($4^{th}$, $6^{th}$, and $8^{th}$ order global accuracy), compact-stencil operators supplied by Fernández [2] performs better both in terms of convergence rate and error magnitude. This is shown in Figure 2 where the solid lines are the various order of accuracy SBP-SAT operators with compact-stencil second derivative operators and the dashed lines indicates the wide-stencil (first derivative twice)
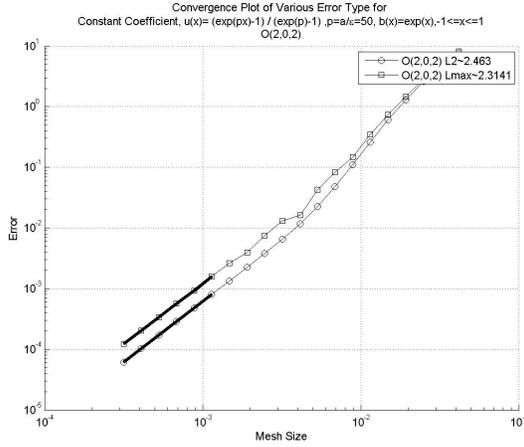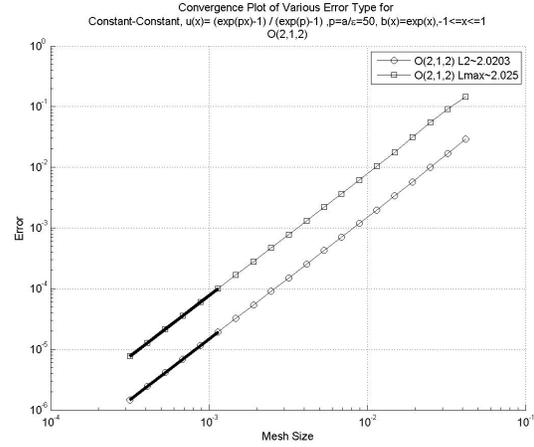
Figure 1: Second Order Operator Convergence Study, Constant-Coefficient O(2,0,2)

SBP-SAT operators. The compact- and wide-stencil operators clearly have a significant difference for the constant-coefficient case. We also note that even at coarse mesh where convergence is in the non-asymptotic region, error for compact-stencil operators are significantly less in magnitude compared to wide-stencil operators.

## 5.2 Variable-Coefficients

We use method of manufactured solution to allow variation of both solution shape and the variable-coefficients. Cases involving sinusoidal function, exponential function, hyperbolic function, and Gaussian functions are examined for both the solution shapes and the variable-coefficients. These four functions are focused because of their unique characteristics near the boundary and relevance in design application [2]. We outline some key results observed for some of these cases.

First we discuss the effect of varying solution shapes and how it affects overall convergence trends in 5.2.1. Next we notice that for some cases, order of accuracy for compact stencil operators collapse to same order of accuracy as wide-stencil operators. We discuss this observation in 5.2.2. Then, we discuss some scenarios where observed convergence rates are higher than theoretical rates. This is presented in 5.2.3. Finally, we compared four operators including an operator that is optimized and discuss the optimization in 5.2.4.
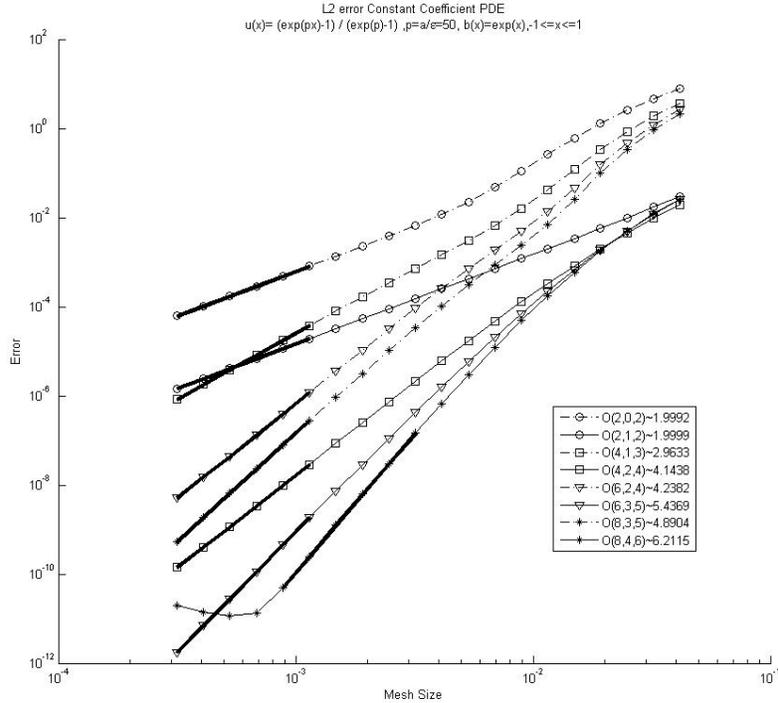
Figure 2: Constant-Coefficient Case Convergence Trend for Various Order of Accuracy

### 5.2.1 Effect of Varying Solution Shapes

We examined the case of only varying shapes of the PDE solutions while fixing the variable-coefficients at exponential function $e^x$. The cases of solution shapes are exponential, hyperbolic cosine, and Gaussian solutions. In particular, hyperbolic cosine function is used to simulate solutions with steep slope near the boundary, and Gaussian function is used to examine effect of steep slope on the interior portion of the mesh. Using operators supplied by Fernández for 4th, 6th, and 8th order interior accuracy with exponential variable-coefficients, we produced convergence trends for combinations of exponential variable-coefficient with the three cases of PDE solutions at various interior order of accuracy.

One particular case with the exponential variable-coefficient $b(x) = e^x$ at very steep hyperbolic cosine solution function $u(x) = cosh(200x)/cosh(200)$ (Figure 4 shows how solution shapes compare for hyperbolic cosine solution shapes used) is presented in Table 1. Corresponding plots are found in Figure 4. We notice conforming trend to the theoretical rate found in works by Fernández and Zingg [2] but the mesh required to reach this rate is much finer than the case of moderately steep function shapes. For example, the convergence trend for a similar setup with the hyperbolic cosine solution function modified to $u(x) = cosh(20x)/cosh(20)$ (Figure 3b
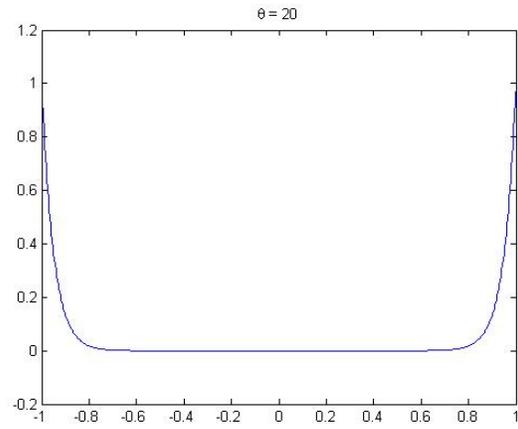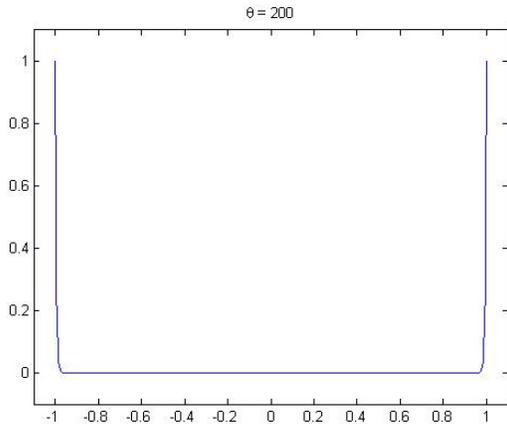
9

(a) Significantly Steep Hyperbolic Cosine Function (b) Moderately Steep Hyperbolic Cosine Function
$u(x) = cosh(200x)/cosh(200)$                 $u(x) = cosh(20x)/cosh(20)$

Figure 3: Hyperbolic Cosine Solution Function Shapes
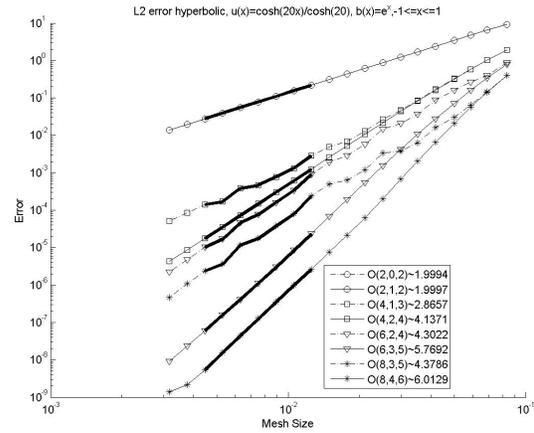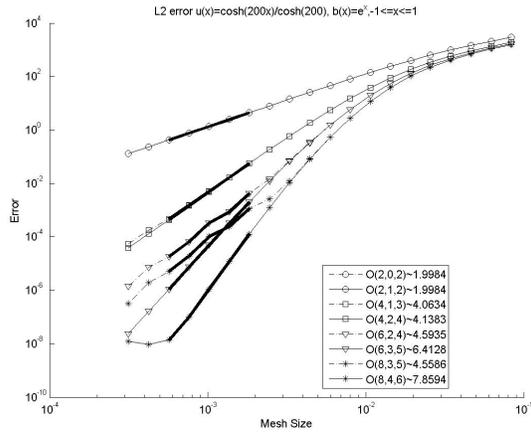


(a) Significantly Steep Hyperbolic Cosine Function (b) Moderately Steep Hyperbolic Cosine Function
$u(x) = cosh(200x)/cosh(200)$                 $u(x) = cosh(20x)/cosh(20)$

Figure 4: Convergence Trend Comparison for Different Shapes of Hyperbolic Cosine Solution Function

shows shape of solution) is shown in Figure 4b. The steep solution case (fig 4a) requires approximately 10 times the mesh grids to achieve same accuracy as the moderately steep case. The initial convergence rate of the steep solution case is quite far from the theoretical value and the solver stays this way for a while until mesh is refined significantly more, before converging to the theoretical rates. This is the non-asymptotic region for which we want to avoid computing the convergence trend on, and can be observed in fig 4a.

10

| solution $u(x)$ | $\|\mathbf{e}\|_2$ Error | Integral Error |
|---|---|---|
| $O(2,0,2)$ | 1.9984 | 1.9983 |
| **O(2,1,2)** | **1.9984** | **1.9983** |
| $O(4,1,3)$ | 4.0634 | 4.1382 |
| **O(4,2,4)** | **4.1322** | **4.1322** |
| $O(6,2,4)$ | 4.5935 | 6.4134 |
| **O(6,3,5)** | **6.4128** | **6.4134** |
| $O(8,3,5)$ | 4.5586 | 8.3692 |
| **O(8,4,6)** | **7.8594** | **7.9807** |

Table 1: Convergence rate for $\|\mathbf{e}\|_2$ and Integral Error for steep gradient hyperbolic cosine solution function case (emphasis shows compact-stencil)

For the rest of experiments in this paper, we will focus on a moderately steep choice of solution to give us the features we want to examine, but at the same time, avoid staying in the non-asymptotic region for too long. For cases involving hyperbolic cosine and exponential function shapes, we found $\frac{cosh(100x)}{cosh(100)}$ and $e^{100x-100}$ to be adequate solution shapes to use.

### 5.2.2 Decreased Compact-Stencil Operator Accuracy

For one particular test scenario, we observe the converging rate for compact-stencil operator collapsing to the same rate for wide-stencil operator of same interior order of accuracy p+1, where normally they should exhibit p+2 accuracy. For example, Figure 5 shows that all the compact-stencil operators except for the second-order operators have convergence rate one below the theoretical rate.

Furthermore, Figure 5 shows that the difference in error magnitude between compact- and wide-stencil operators have almost vanished in some case. In most other cases, the prevailing trend observed was that the compact-stencil operators would have much less error compared to the wide-stencil operators.

Also from Figure 5, the fact that eighth order operators have hit the machine accuracy level and still exhibit this behavior implies that this phenomena is occurring in asymptotic region.

### 5.2.3 Increased Operator Accuracy

We observed that in some cases, operator can outperform their theoretical accuracy. In particular, they occur when the solution shape involves Gaussian, in which case both wide- and
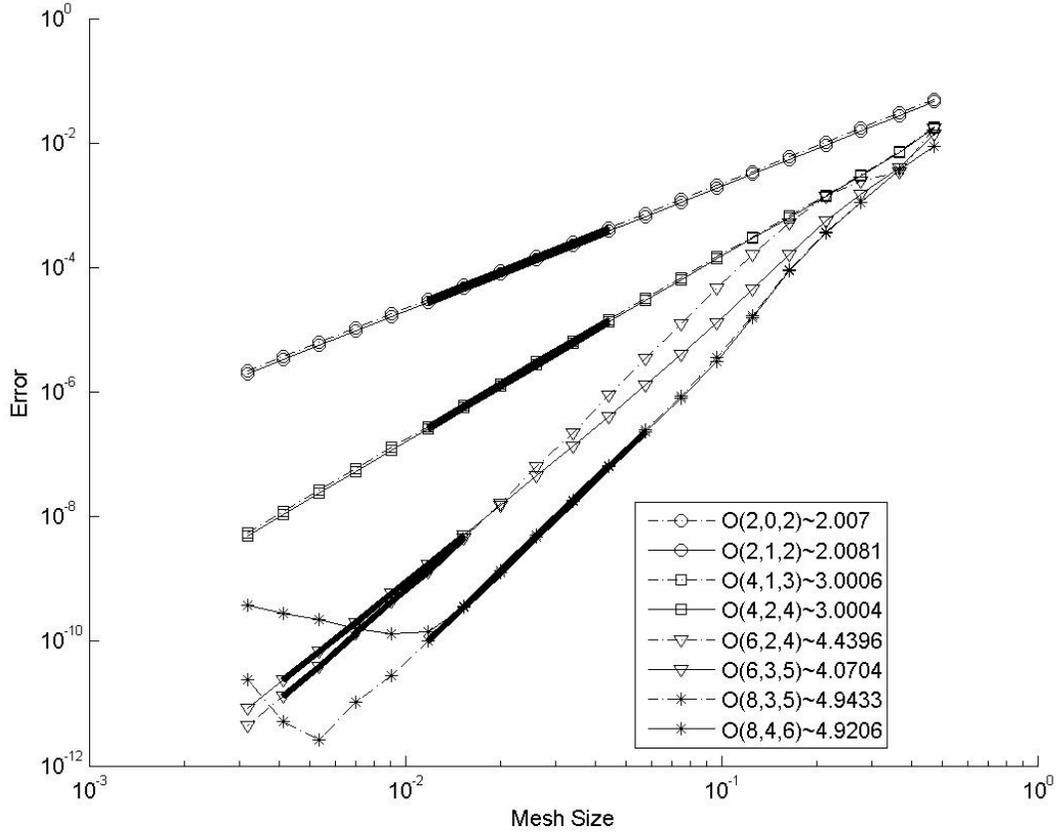
Figure 5: p+1 Order Accuracy Compact-Stencil Operators, $u(x) = sin(x), b(x) = e^{-x^2}$

compact-stencil operators exhibit 2p order of accuracy. Another case in which this occur is when wide-stencil operators exhibit p+2 order of accuracy, which is same as the theoretical order of accuracy for compact stencil operators.

The first case is observed for most of the results involving Gaussian solution shape, which is not surprising. This is because the function is nearly zero for the Gaussian solution shape cases, and from theoretical accuracy of SBP operators for second derivatives, we know that the primary difference in order of accuracy between compact- and wide-stencil operator lies in their boundary order accuracy of p+1 and p+2 respectively. Therefore, with the function taking such small values near boundaries, the Gaussian solution shape might be showing exceptional operator accuracy because the interior error is actually dominating. For example, Figure 6 shows a particular Gaussian solution shape scenario. The corresponding compact- and wide-stencil operators seem to exhibit theoretical order of accuracy for the interior blocks which are 2p.

More interestingly we have the second case where increased order of accuracy was observed for

12
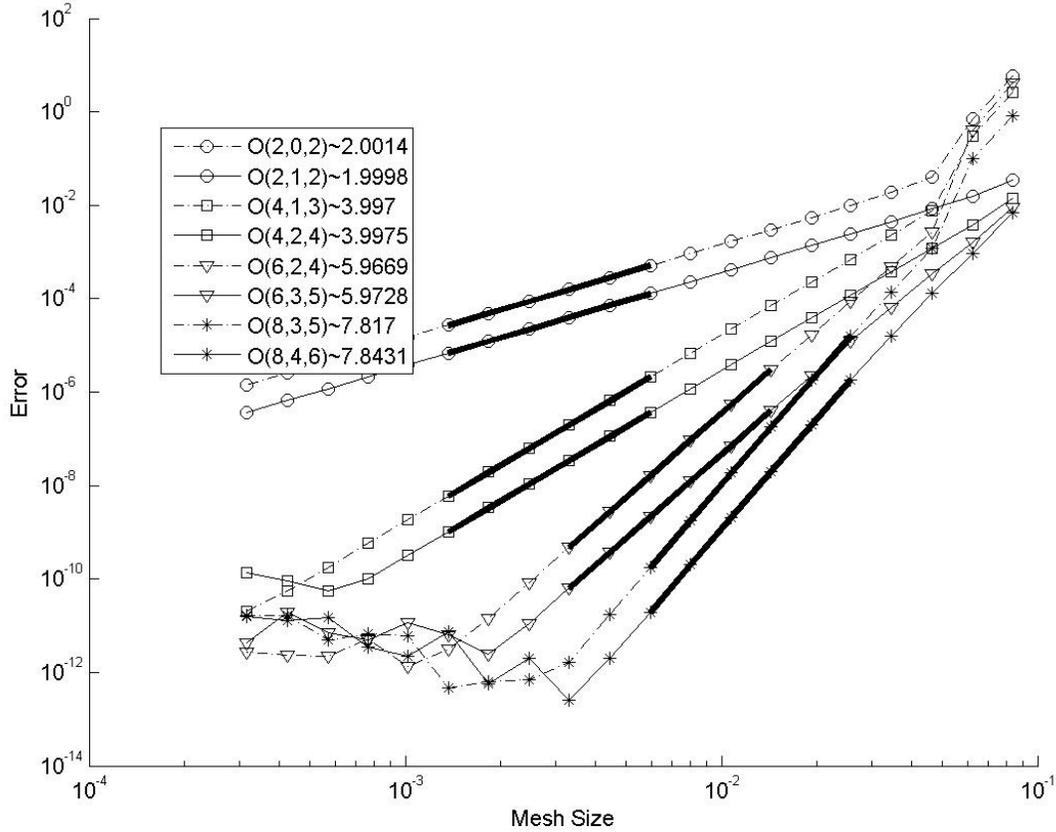
Figure 6: $\|\mathbf{e}\|_2$ norm solution error, $u(x) = exp(-100x^2), b(x) = e^{-x^2}$

non-Gaussian solution shapes. For example, convergence trend of a case involving sinusoidal solution shape and hyperbolic cosine variable-coefficient is shown in Figure 7. The solution function does not smoothly converge to zero unlike the Gaussian case near the boundary, which is at $-2\pi$ and $2\pi$ covering two periods. There is also an important distinction between this case and the Gaussian cases in the improvement pattern. Instead of going to the interior order of accuracy 2p, the phenomenon observed involves the wide-stencil operator improving to the order of accuracy of the corresponding compact-stencil operator at p+2 for operators except the eight order wide-stencil operator (O(8,3,5)). The compact-stencil operators remain at their theoretical order of accuracy p+2. Furthermore, the error magnitude at coarser mesh for compact-and wide-stencil operator are very similar unlike the other cases where we observed significantly lower error magnitude for compact-stencil operators even at very coarse mesh.
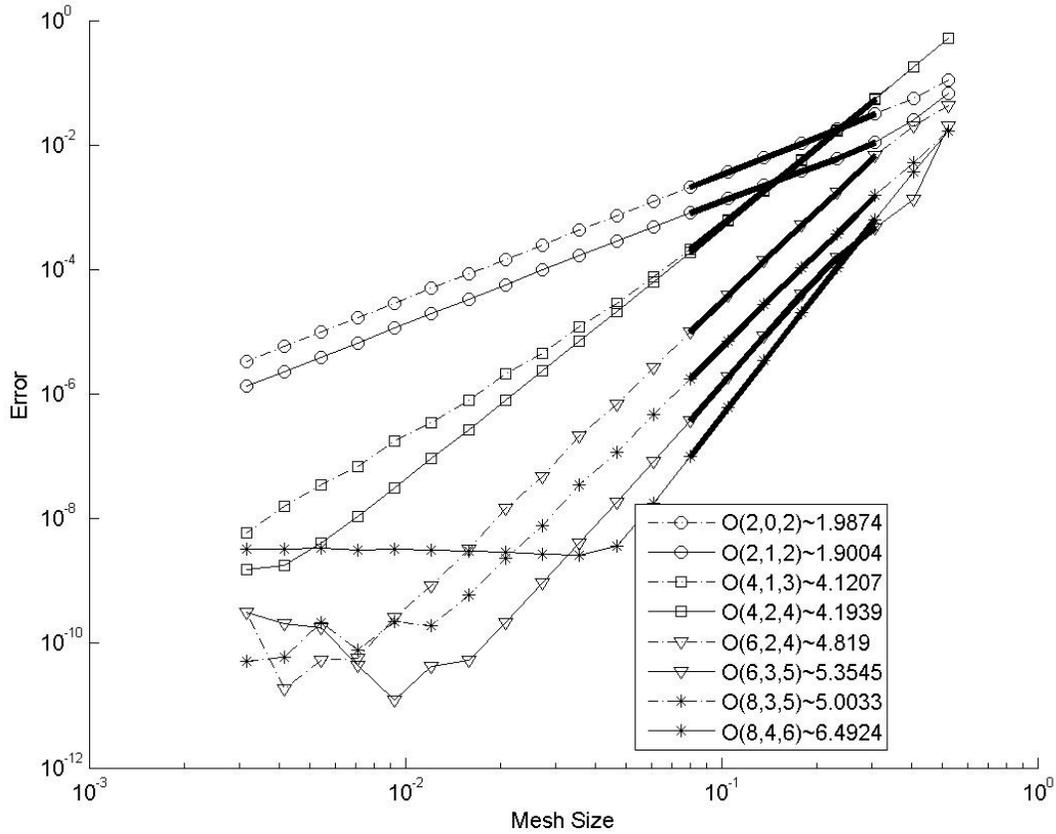
Figure 7: $\|\mathbf{e}\|_2$ norm solution error, $u(x) = sin(x), b = cosh(x)$

### 5.2.4 Varying Operator

In addition to examining the general purpose operator by Fernández against the wide-stencil operator, we also examined how Mattsson's operator and an optimized operator performed. The optimized operator is fourth order compact-stencil operator also generated by David Fernández at University of Toronto Institute for Aerospace Studies.

The result for fourth order operator is analyzed in Figure 8. We can see a significant improvement for the optimized operator. In fact, the error during the asymptotic region for the operator was only **33.7%** of the original error. This is a significant reduction, showing that SBP-SAT operators can be optimized to give improvement in performance.

Further experiments on the optimized operator indicate that the improvement is heavily
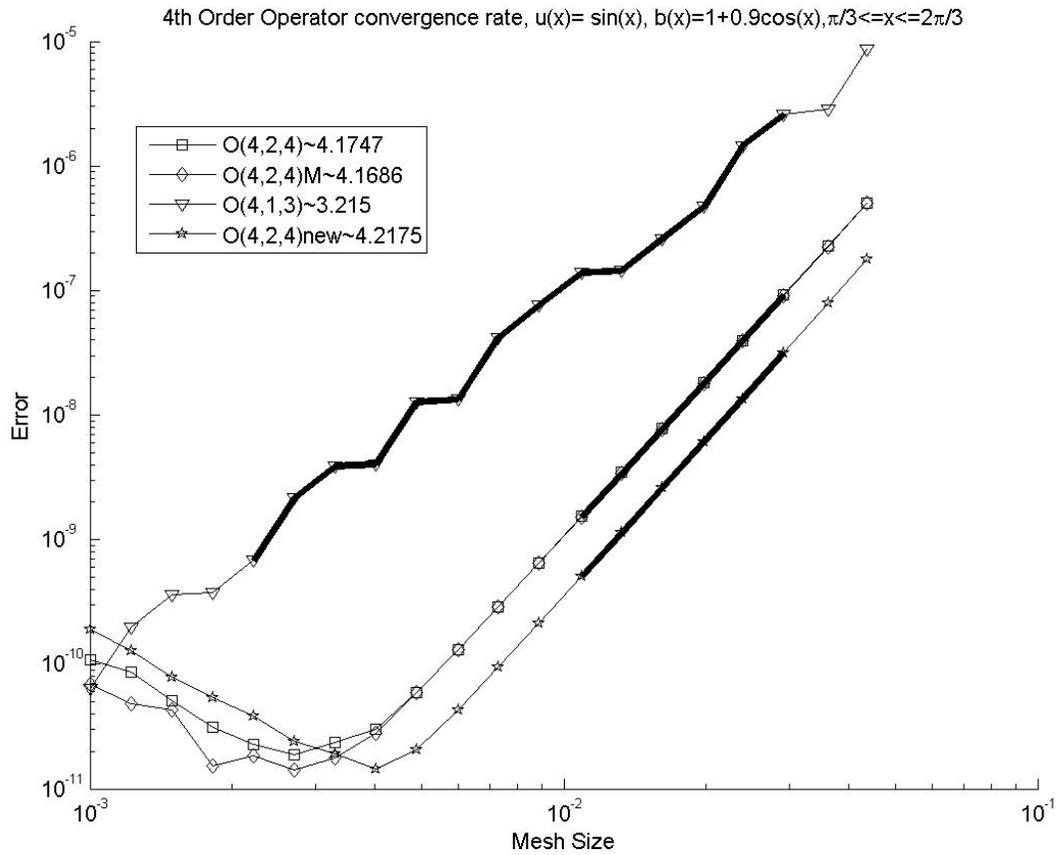
Figure 8: $\|\mathbf{e}\|_2$ norm solution error fourth order operators, $u(x) = sin(x), b = 1 + 0.9cos(x)$

dependent on **solution shape**. Looking at other combination of this solution shape with other variable-coefficient, we observe the improvement is consistent. However, when we move to other solution shape, for example exponential solution shape with exponential variable-coefficient in Figure 9, the improvement is no longer persistent.
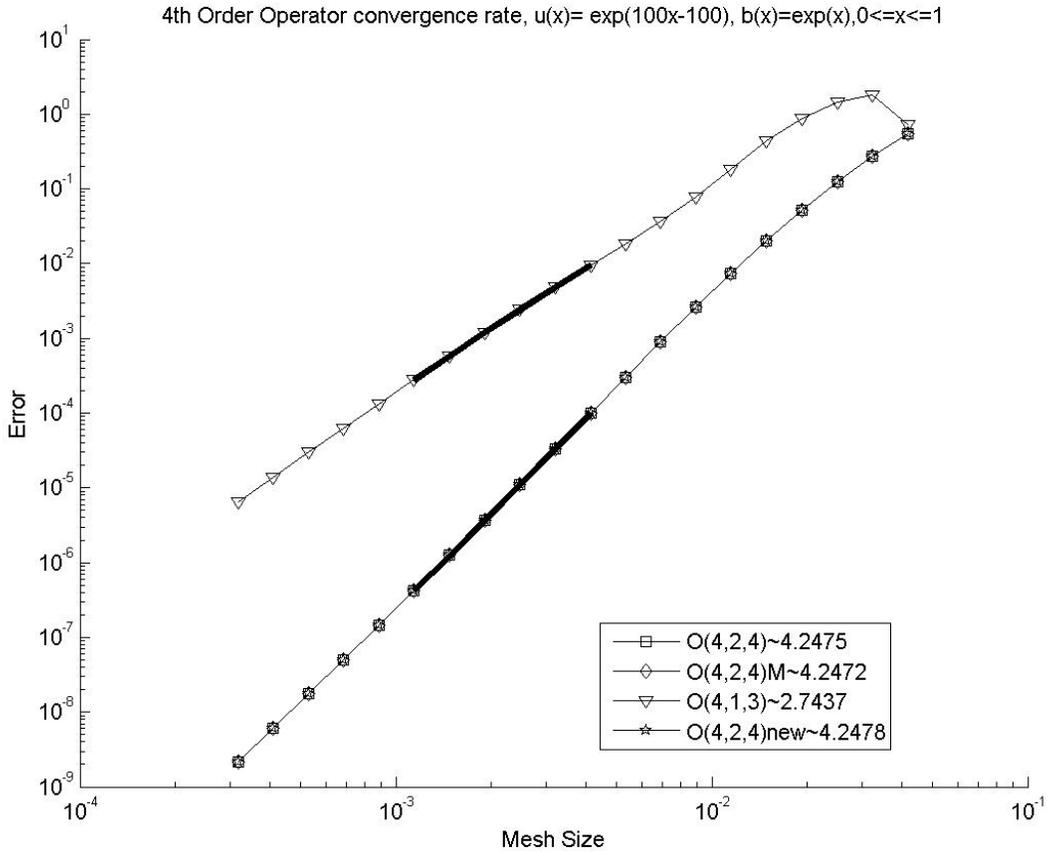
Figure 9: $\|\mathbf{e}\|_2$ norm solution error fourth order operators, $u(x) = e^{100x-100}, b = e^x$

# 6  Conclusion

This thesis provides insight into the performance of summation-by-parts (SBP) operators for second-derivative with variable-coefficients using simultaneous-approximation-terms (SAT). By characterizing the accuracy, we are more knowledgeable in how the SBP discretized second-derivative operators will perform for different problems. The thesis shows that optimization of the SBP-SAT method can provide significant error reduction. This means coarser mesh can be used for optimized SBP operators which leads to reduction in computational cost. The knowledge gained from the results of this thesis can contribute to development of more efficient discretization of second-derivative operator involving the SBP method by outlining some important problem-specific observations.

There are still some more parts of the operator that remain to be characterized. We have not had a chance to examine the SBP-SAT operator free parameters which comes from solving the

system of equations involving SBP property. In the interim report I have briefly explored the optimization parameter for second order SBP-SAT operator, particularly the $c33$ operator. How SBP-SAT operator parameter choice relates to the various choices of coefficients and solution shapes remains to be explored.

# References

[1] Mark H. Carpenter, David Gottlieb, and Saul Abarbanel. Time-Stable boundary conditions for finite-difference schemes solving hyperbolic systems: methodology and application to higher-order compact schemes. *AIAA*, 2012.

[2] David C. Del Rey Fernández and David W. Zingg. High-order compact-stencil summation-by-parts operators for the second derivative with variable coefficients. *ICCFD7*, 2012.

[3] Thomas H. Pulliam Harvard Lomax and David W. Zingg. *Fundamentals of Computational Fluid Dynamics*. Springer, 2004.

[4] Jason E. Hicken and David W. Zingg. Parallel Newton-Krylov solver for the Euler equations discretized using Simultaneous-Approximation Terms. *AIAA*, 46(11):2773–2786, 2008.

[5] Heinz-Otto Kreiss and Joseph Oliger. Comparison of accurate methods for the integration of hyperbolic equations. *Telus*, 24(3):199–215, 1972.

[6] Osusky. L. and Zingg. D. W. A parallel Newtion-Krylov-Shur flow solver for the Reynolds-Averaged Navier-Stokes equations. *AIAA*, 2012.

[7] Ken Mattsson. Summation by parts operators for finite difference approximations of second-derivatives with variable coefficients. *SIAM J. Sci. Comput.*, 51(3):650–682, 2012.

[8] Ken Mattsson and Jan Nordström. Summation by parts operators for finite difference approximations of second-derivatives. *Journal of Computational Physics*, 199(2):503–540, 2004.

[9] M. Nemec and D. W. Zingg. Newton-Krylov algorithm for aerodyanmic design using the Navier-Stokes equations. *AIAA Journal*, 40(6):1146–1154, 2002.

[10] Jan Nordström and Mark H. Carpenter. Higher-order finite difference methods, multidimensional linear problems, and curvilinear coordinates. *Journal of Computational Physics*, 173(1):149–174, 2001.

[11] Bo Strand. Summation by parts for finite difference approximations for d/dx. *Journal of Computational Physics*, 110(1):47–67, 1994.

[12] David W. Zingg. Comparison of high-accuracy finite-difference methods for linear wave propagation. *SIAM J. Sci. Comput.*, 22(2):476–502, 2012.

# Appendix A   Observed Order of Accuracy Table

| u(x) | b(x) | order | 2 − norm − error | inf − norm − error |
|------|------|-------|------------------|--------------------|
| sin(x) | 1 | $O(2,0,2)$ | 2.0031 | 2.0028 |
| | | $O(2,1,2)$ | 2.0020 | 2.0003 |
| | | $O(4,1,3)$ | 2.9975 | 2.9790 |
| | | $O(4,2,4)$ | 4.1956 | 4.2323 |
| | | $O(6,2,4)$ | 4.8275 | 4.8481 |
| | | $O(6,3,5)$ | 5.4336 | 4.9562 |
| | | $O(8,3,5)$ | 4.9613 | 4.9316 |
| | | $O(8,4,6)$ | 7.0033 | 6.6294 |

| u(x) | b(x) | order | 2 − norm − error | inf − norm − error |
|------|------|-------|------------------|--------------------|
| sin(10x) | 1 | $O(2,0,2)$ | 2.0003 | 2.0003 |
| | | $O(2,1,2)$ | 2.0003 | 2.0002 |
| | | $O(4,1,3)$ | 3.1954 | 3.0422 |
| | | $O(4,2,4)$ | 4.0483 | 4.0534 |
| | | $O(6,2,4)$ | 5.1488 | 4.9575 |
| | | $O(6,3,5)$ | 5.4855 | 4.9907 |
| | | $O(8,3,5)$ | 5.1242 | 4.9229 |
| | | $O(8,4,6)$ | 7.3525 | 6.7715 |

| u(x) | b(x) | order | 2 − norm − error | inf − norm − error |
|------|------|-------|------------------|--------------------|
| exp(10x-10) | 1 | $O(2,0,2)$ | 2.0002 | 1.9791 |
| | | $O(2,1,2)$ | 2.0014 | 1.9999 |
| | | $O(4,1,3)$ | 2.9319 | 2.9208 |
| | | $O(4,2,4)$ | 4.0826 | 4.0829 |
| | | $O(6,2,4)$ | 4.3983 | 3.9727 |
| | | $O(6,3,5)$ | 5.3510 | 4.8569 |
| | | $O(8,3,5)$ | 4.7830 | 4.7742 |
| | | $O(8,4,6)$ | 6.0344 | 5.5338 |

| u(x) | b(x) | order | 2 − norm − error | inf − norm − error |
|------|------|-------|------------------|--------------------|
| exp(100x-100) | 1 | $O(2,0,2)$ | 1.9997 | 1.9943 |
| | | $O(2,1,2)$ | 1.9997 | 1.9995 |

| u(x) | b(x) | order | $2-norm-error$ | $inf-norm-error$ |
|---|---|---|---|---|
| | | $O(4,1,3)$ | 3.8155 | 3.5946 |
| | | $O(4,2,4)$ | 4.1137 | 4.1139 |
| | | $O(6,2,4)$ | 4.3524 | 3.9451 |
| | | $O(6,3,5)$ | 6.3593 | 5.1586 |
| | | $O(8,3,5)$ | 4.7095 | 4.6993 |
| | | $O(8,4,6)$ | 6.8019 | 5.4556 |
| **u(x)** | **b(x)** | *order* | $2-norm-error$ | $inf-norm-error$ |
| exp(-100x^2) | 1 | $O(2,0,2)$ | 2.1410 | 2.1477 |
| | | $O(2,1,2)$ | 2.0263 | 2.0056 |
| | | $O(4,1,3)$ | 3.9498 | 3.9180 |
| | | $O(4,2,4)$ | 3.9154 | 3.8658 |
| | | $O(6,2,4)$ | 5.7584 | 5.6993 |
| | | $O(6,3,5)$ | 5.7657 | 5.6896 |
| | | $O(8,3,5)$ | 7.5139 | 7.4264 |
| | | $O(8,4,6)$ | 7.5557 | 7.4543 |
| **u(x)** | **b(x)** | *order* | $2-norm-error$ | $inf-norm-error$ |
| cosh(10x)/cosh(10) | 1 | $O(2,0,2)$ | 1.9953 | 1.9859 |
| | | $O(2,1,2)$ | 2.0004 | 1.9998 |
| | | $O(4,1,3)$ | 2.7788 | 2.8657 |
| | | $O(4,2,4)$ | 4.1036 | 4.1032 |
| | | $O(6,2,4)$ | 4.2837 | 3.9583 |
| | | $O(6,3,5)$ | 5.2986 | 4.7864 |
| | | $O(8,3,5)$ | 4.2708 | 4.7660 |
| | | $O(8,4,6)$ | 6.0529 | 5.5353 |
| **u(x)** | **b(x)** | *order* | $2-norm-error$ | $inf-norm-error$ |
| sin(x) | cosh(x) | $O(2,0,2)$ | 1.9874 | 2.0124 |
| | | $O(2,1,2)$ | 1.9004 | 1.8580 |
| | | $O(4,1,3)$ | 4.1207 | 3.9944 |
| | | $O(4,2,4)$ | 4.1939 | 4.1913 |
| | | $O(6,2,4)$ | 4.8190 | 4.7231 |
| | | $O(6,3,5)$ | 5.3545 | 4.6581 |
| | | $O(8,3,5)$ | 5.0033 | 5.0191 |

| | | $O(8,4,6)$ | 6.4924 | 6.0631 |
|---|---|---|---|---|
| **u(x)** | **b(x)** | *order* | $2-norm-error$ | $inf-norm-error$ |
| | | $O(2,0,2)$ | 2.0070 | 2.0745 |
| | | $O(2,1,2)$ | 2.0081 | 2.0000 |
| | | $O(4,1,3)$ | 3.0006 | 2.9861 |
| | | $O(4,2,4)$ | 3.0004 | 2.9861 |
| sin(x) | exp(-x^2) | $O(6,2,4)$ | 4.4396 | 4.3533 |
| | | $O(6,3,5)$ | 4.0704 | 4.0081 |
| | | $O(8,3,5)$ | 4.9433 | 4.9347 |
| | | $O(8,4,6)$ | 4.9206 | 4.9351 |
| **u(x)** | **b(x)** | *order* | $2-norm-error$ | $inf-norm-error$ |
| | | $O(2,0,2)$ | 2.0010 | 2.0028 |
| | | $O(2,1,2)$ | 2.0011 | 2.0028 |
| | | $O(4,1,3)$ | 4.1976 | 4.1993 |
| | | $O(4,2,4)$ | 4.1976 | 4.1997 |
| sin(x) | exp(x) | $O(6,2,4)$ | 7.3687 | 7.2580 |
| | | $O(6,3,5)$ | 7.3720 | 7.3675 |
| | | $O(8,3,5)$ | 8.2075 | 8.0159 |
| | | $O(8,4,6)$ | 8.3575 | 8.2821 |
| **u(x)** | **b(x)** | *order* | $2-norm-error$ | $inf-norm-error$ |
| | | $O(2,0,2)$ | 2.1726 | 2.0546 |
| | | $O(2,1,2)$ | 2.0032 | 2.0002 |
| | | $O(4,1,3)$ | 3.8944 | 3.0693 |
| | | $O(4,2,4)$ | 4.2008 | 3.9092 |
| sin(x) | cos(x) | $O(6,2,4)$ | 4.4607 | 4.0385 |
| | | $O(6,3,5)$ | 5.5737 | 5.0615 |
| | | $O(8,3,5)$ | 4.9077 | 5.0582 |
| | | $O(8,4,6)$ | 6.0632 | 5.4462 |
| **u(x)** | **b(x)** | *order* | $2-norm-error$ | $inf-norm-error$ |
| | | $O(2,0,2)$ | 2.0020 | 1.9785 |
| | | $O(2,1,2)$ | 2.0002 | 2.0000 |
| | | $O(4,1,3)$ | 2.9838 | 2.9929 |

exp(10x-10)        exp(x)

| u(x) | b(x) | order | 2 − norm − error | Inf − norm − error |
|---|---|---|---|---|
| | | $O(4,2,4)$ | 4.1702 | 3.9165 |
| | | $O(6,2,4)$ | 4.2324 | 4.0052 |
| | | $O(6,3,5)$ | 5.3243 | 4.9154 |
| | | $O(8,3,5)$ | 4.8844 | 4.8674 |
| | | $O(8,4,6)$ | 6.1261 | 5.5660 |

| u(x) | b(x) | order | 2 − norm − error | Inf − norm − error |
|---|---|---|---|---|
| | | $O(2,0,2)$ | 2.0359 | 1.8045 |
| | | $O(2,1,2)$ | 2.0246 | 1.9980 |
| | | $O(4,1,3)$ | 2.7437 | 2.7631 |
| exp(100x-100) | exp(x) | $O(4,2,4)$ | 4.2475 | 4.1358 |
| | | $O(6,2,4)$ | 4.1423 | 3.7644 |
| | | $O(6,3,5)$ | 4.9420 | 4.4443 |
| | | $O(8,3,5)$ | 4.2417 | 4.2511 |
| | | $O(8,4,6)$ | 5.7332 | 5.2274 |

| u(x) | b(x) | order | 2 − norm − error | inf − norm − error |
|---|---|---|---|---|
| | | $O(2,0,2)$ | 1.9957 | 1.9888 |
| | | $O(2,1,2)$ | 1.9958 | 1.9949 |
| | | $O(4,1,3)$ | 3.9247 | 3.8632 |
| cosh(100x)/cosh(100) | cosh(x) | $O(4,2,4)$ | 4.0962 | 4.0951 |
| | | $O(6,2,4)$ | 4.2418 | 3.7463 |
| | | $O(6,3,5)$ | 6.1663 | 5.6763 |
| | | $O(8,3,5)$ | 4.0435 | 4.3037 |
| | | $O(8,4,6)$ | 7.6450 | 6.3762 |

| u(x) | b(x) | order | 2 − norm − error | inf − norm − error |
|---|---|---|---|---|
| | | $O(2,0,2)$ | 2.0014 | 2.0024 |
| | | $O(2,1,2)$ | 1.9998 | 2.0003 |
| | | $O(4,1,3)$ | 3.9970 | 3.9970 |
| exp(-100x^2) | exp(-x^2) | $O(4,2,4)$ | 3.9975 | 3.9983 |
| | | $O(6,2,4)$ | 5.9669 | 5.9616 |
| | | $O(6,3,5)$ | 5.9728 | 5.9689 |
| | | $O(8,3,5)$ | 7.8170 | 7.8166 |
| | | $O(8,4,6)$ | 7.8431 | 7.8410 |

# Appendix B    $\|\mathbf{e}\|_2$ Convergence Plots
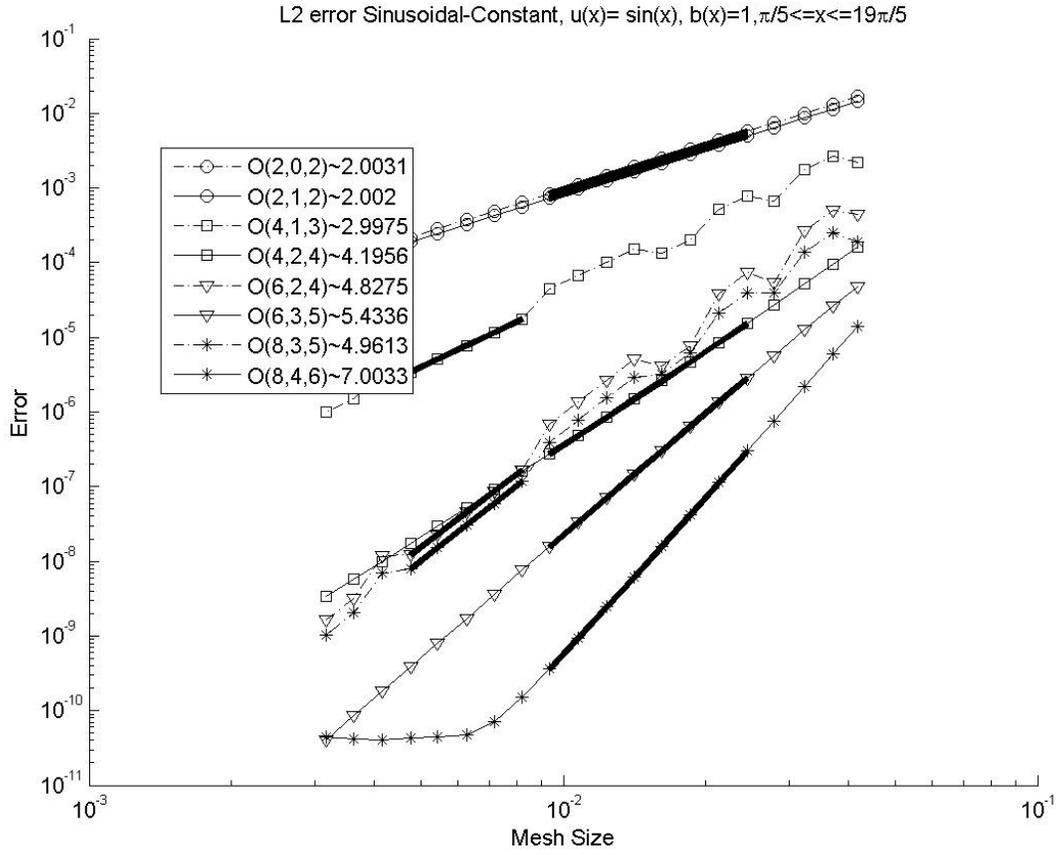


Figure 10: $\|\mathbf{e}\|_2$ Convergence Plot, $u(x) = sin(x), b(x) = 1$
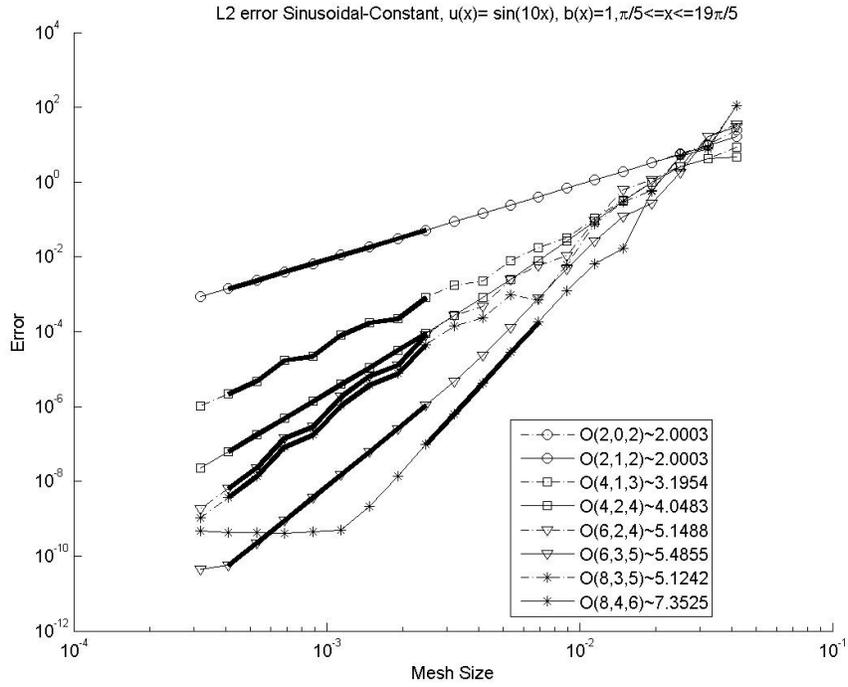
Figure 11: $\|\mathbf{e}\|_2$ Convergence Plot, $u(x) = sin(10x), b(x) = 1$
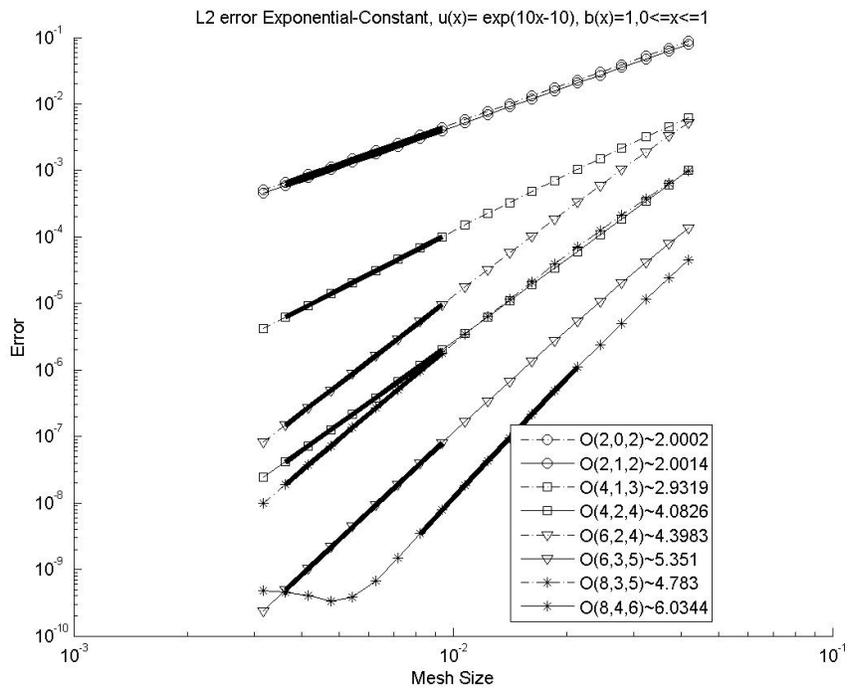


Figure 12: $\|\mathbf{e}\|_2$ Convergence Plot, $u(x) = e^{10x-10}, b(x) = 1$
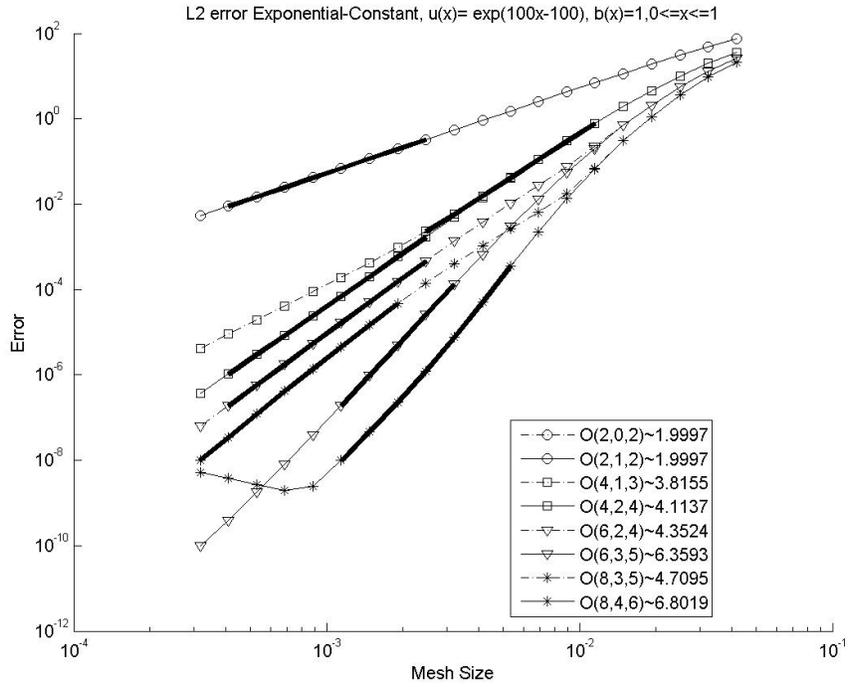
Figure 13: $\|\mathbf{e}\|_2$ Convergence Plot, $u(x) = e^{100x-100}, b(x) = 1$
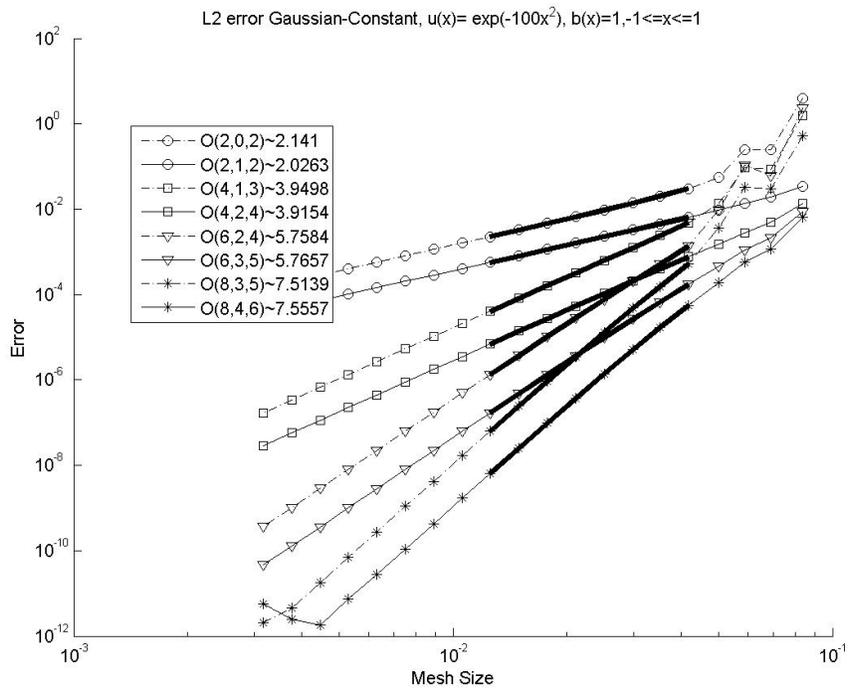


Figure 14: $\|\mathbf{e}\|_2$ Convergence Plot, $u(x) = e^{-100x^2}, b(x) = 1$

25
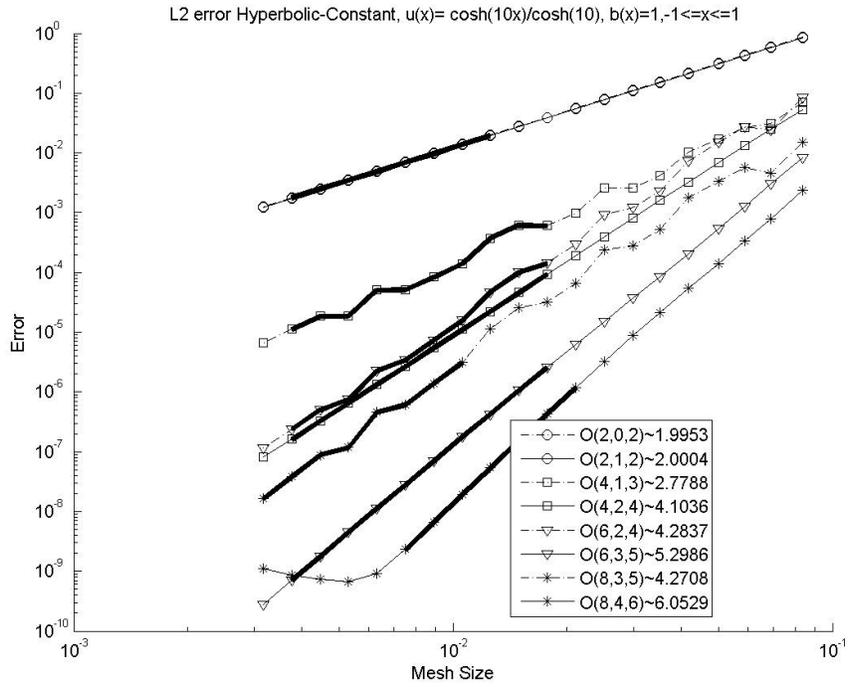
Figure 15: $\|\mathbf{e}\|_2$ Convergence Plot, $u(x) = cosh(10x)/cosh(10), b(x) = 1$
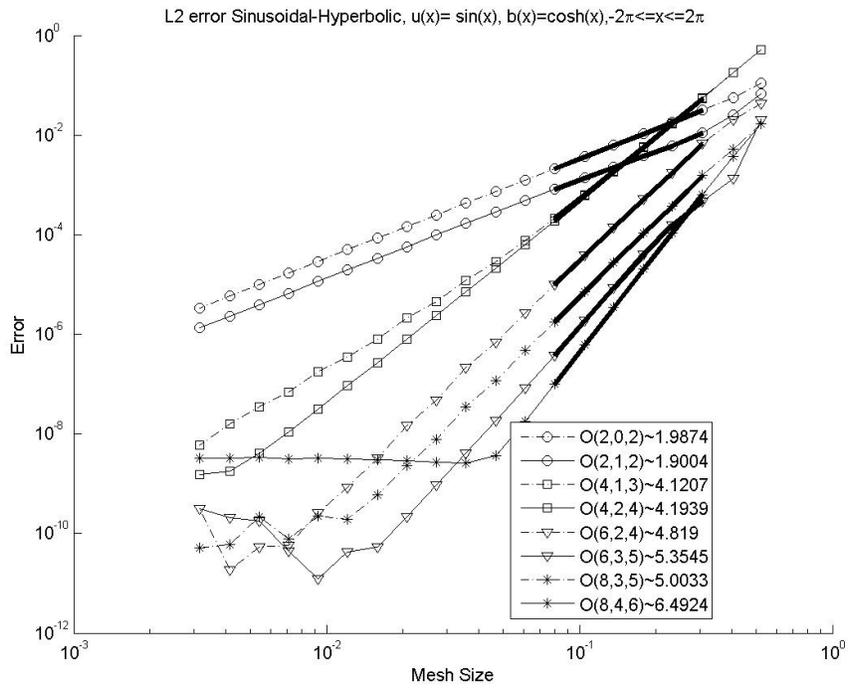


Figure 16: $\|\mathbf{e}\|_2$ Convergence Plot, $u(x) = sin(x), b(x) = cosh(x)$
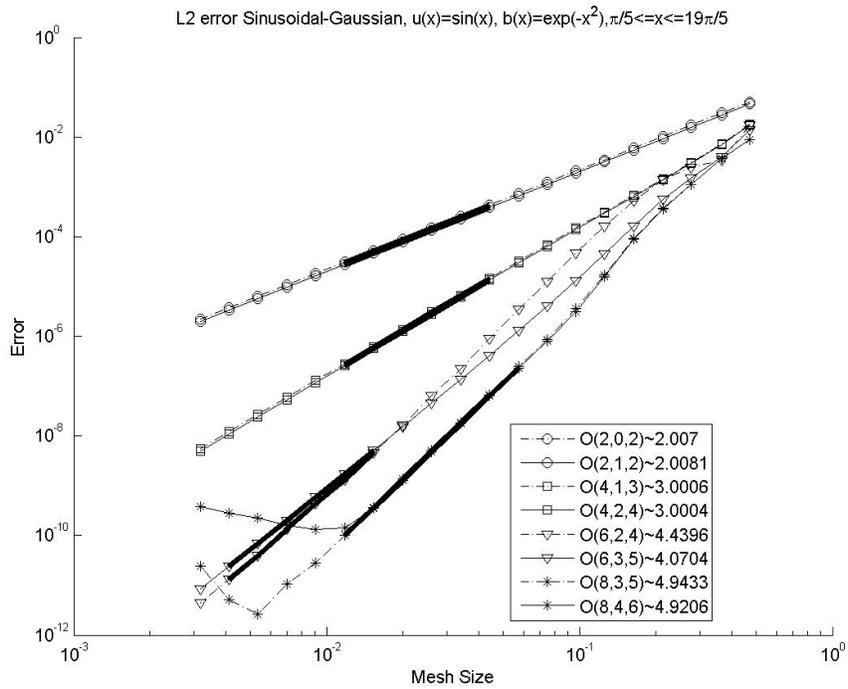
Figure 17: $\|\mathbf{e}\|_2$ Convergence Plot, $u(x) = sin(x), b(x) = e^{-x^2}$
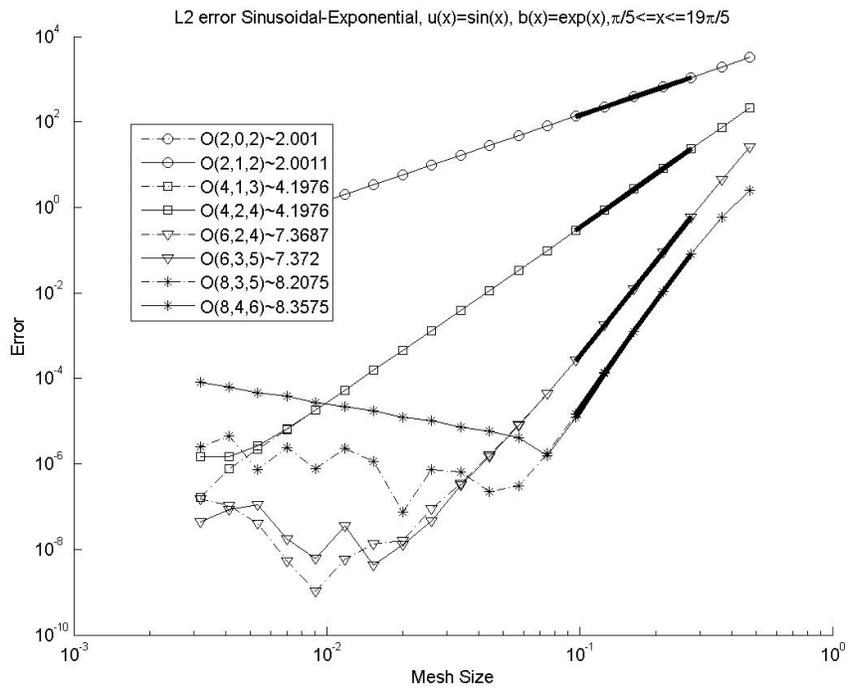


Figure 18: $\|\mathbf{e}\|_2$ Convergence Plot, $u(x) = sin(x), b(x) = e^x$
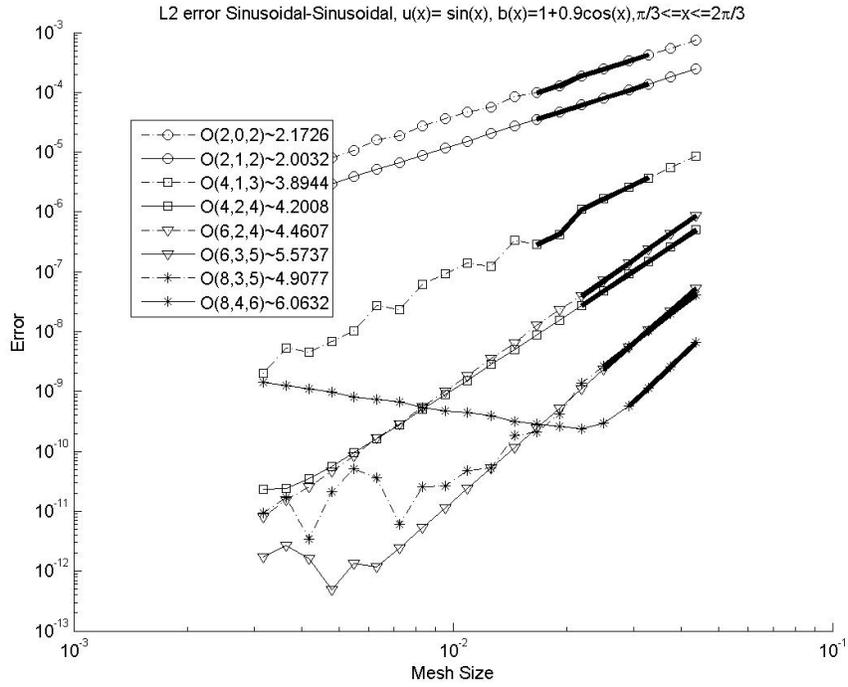
Figure 19: $\|\mathbf{e}\|_2$ Convergence Plot, $u(x) = sin(x), b(x) = cos(x)$
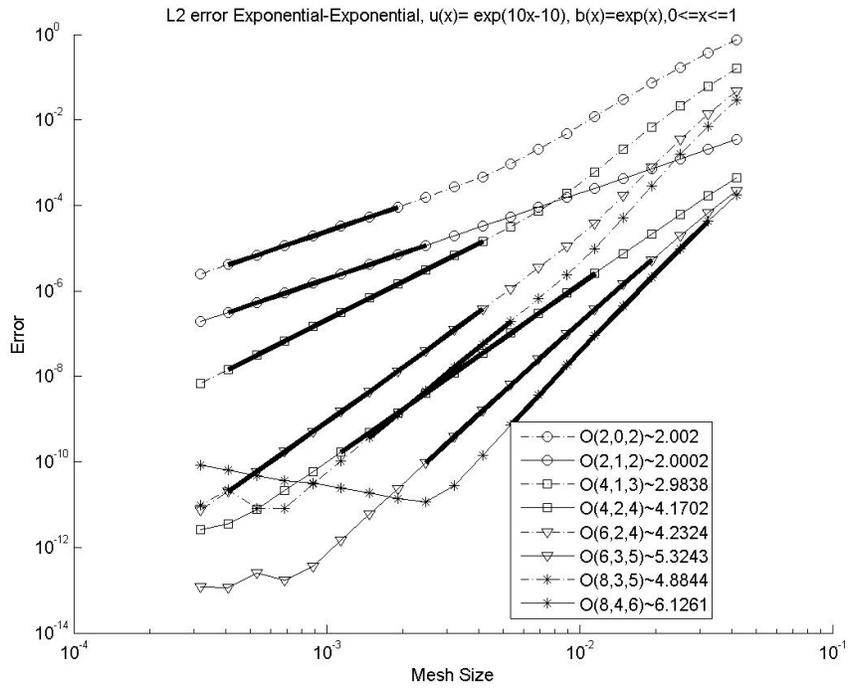


Figure 20: $\|\mathbf{e}\|_2$ Convergence Plot, $u(x) = e^{10x-10}, b(x) = e^x$

28
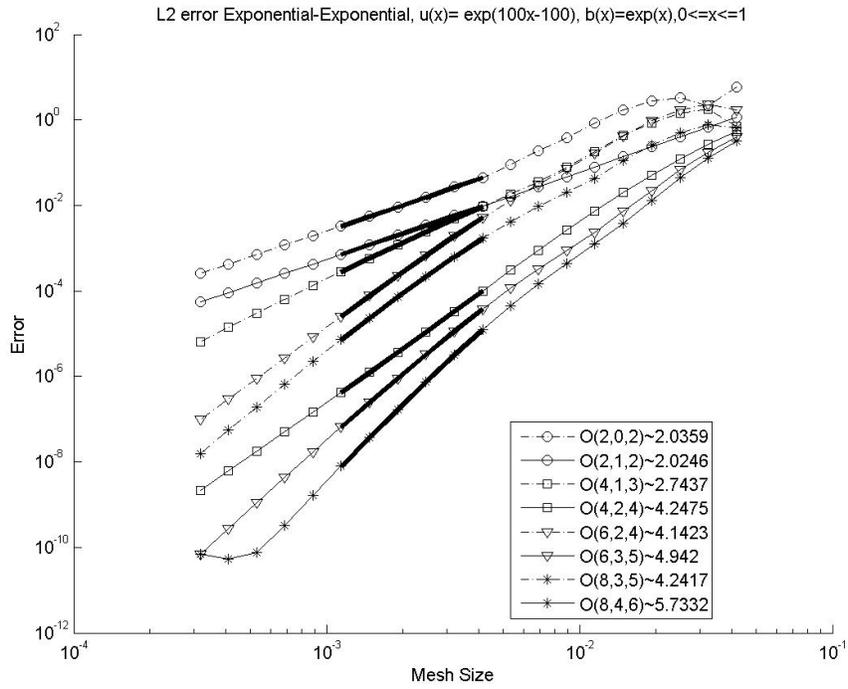
Figure 21: $\|\mathbf{e}\|_2$ Convergence Plot, $u(x) = e^{100x-100}, b(x) = e^x$
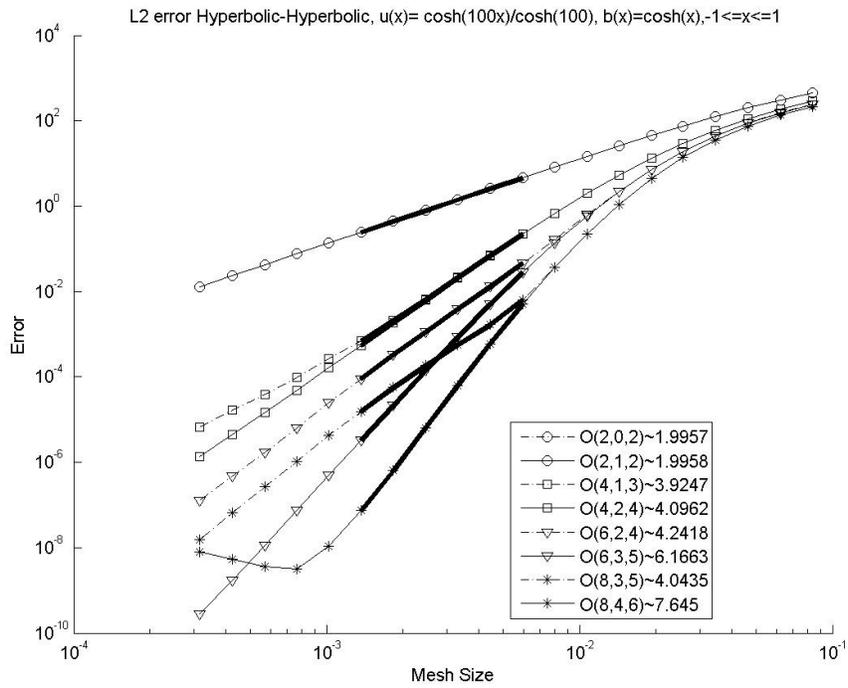


Figure 22: $\|\mathbf{e}\|_2$ Convergence Plot, $u(x) = cosh(100x)/cosh(100), b(x) = cosh(x)$
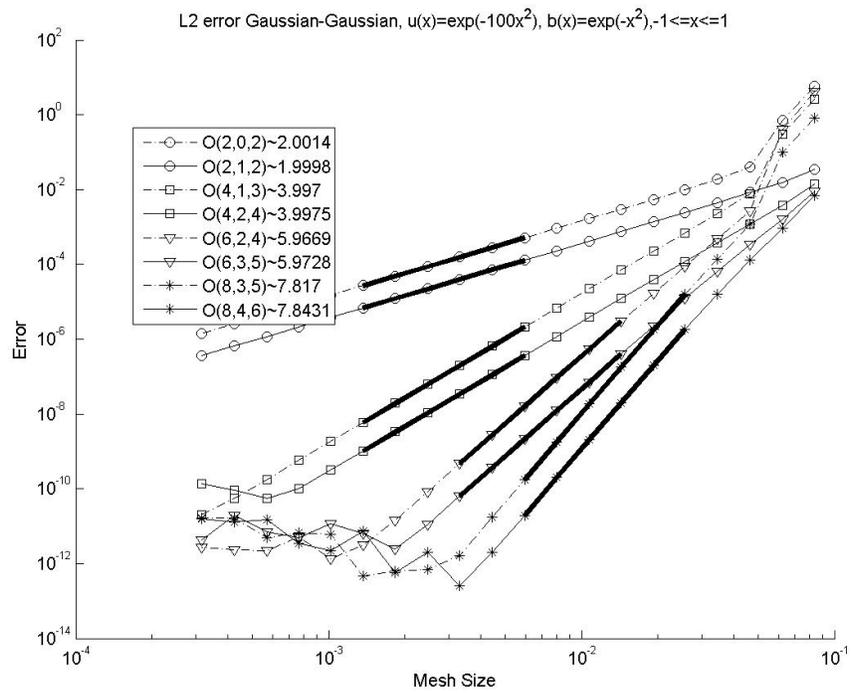
Figure 23: $\|\mathbf{e}\|_2$ Convergence Plot, $u(x) = e^{-100x^2}, b(x) = e^{-x^2}$