Error Estimates and Communication Overhead in the Computation of the Bidomain Equations on the Distributed Memory Parallel Blue Gene/L Supercomputer

M Reumann[1], BG Fitch[1], A Rayshubskiy[1], DL Weiss[2], G Seemann[2], O Doessel[2], MC Pitman[1], JJ Rice[1]

[1]IBM T J Watson Research Center, Yorktown Heights, NY, USA
[2]Institute of Biomedical Engineering, Karlsruhe Institute of Technology, Karlsruhe, Germany

Abstract

Increasing biophysical detail in multi physical, multiscale cardiac model will demand higher levels of parallelism in multi-core approaches to obtain fast simulation times. As an example of such a highly parallel multi-core approaches, we develop a completely distributed bidomain cardiac model implemented on the IBM Blue Gene/L architecture. A tissue block of size 50 x 50 x 100 cubic elements based on ten Tusscher et al. (2004) cell model is distributed on 512 computational nodes. The extracellular potential is calculated by the Gauss-Seidel (GS) iterative method that typically requires high levels of inter-processor communication. Specifically, the GS method requires knowledge of all cellular potentials at each of it iterative step. In the absence of shared memory, the values are communicated with substantial overhead. We attempted to reduce communication overhead by computing the extracellular potential only every 5th time step for the integration of the cell models. We also investigated the effects of reducing inter-processor communication to every 5th, 10th, 50th iteration or no communication within the GS iteration. While technically incorrect, these approximation had little impact on numerical convergence or accuracy for the simulations tested. The results suggest some heuristic approaches may further reduce the inter-processor communication to improve the execution time of large-scale simulations.

1. Introduction

While CPU power has followed Moore's Law since the 1970s, the consensus is that clock speeds will not appreciable increase in the foreseeable future with current silicon technology. In light of this, further increases in computation power will demand higher levels of parallelism in multi-core approaches. In tissue-level cardiac models, the parallelism will require distribution of workloads, generally based on volumetric decomposition of the heart tissue itself. The balancing of computational load and inter-processor communication become paramount in a highly distributed environment where memory is not shared between all processors (typical in non-SMP machines such as clusters). We present such an example with bidomain tissue-level cardiac models implemented on the Blue Gene/L architecture. Standard bidomain calculations [1] require knowledge of the electrical potential at all points in the 3D mesh, and hence, the results demonstrate an important use case in a completely distributed environment that does not permit data coalescence into a commonly-shared memory space that is typical in existing implementations.

2. Methods

We model a tissue block of size 50 x 50 x 100 cubic elements is used a test case for the algorithm development and testing. Computation is distributed on 512 computational nodes of an IBM Blue Gene/L supercomputer with distributed memory. A single stimulus is set to create a propagating wave front per run. The ten Tusscher et al. cell model [2] and the bidomain equations are solved for heterogeneous anisotropic tissue by the Saleheen et al. [3, 4] formulation.

The extracellular potential is calculated by the Gauss-Seidel (GS) iterative method using either the global sum of squared residuals (gSSR) across all computational nodes with gSSR < 10^{-6} as termination criterion or if a maximum number of iterations is reached. The gSSR is caluculated as shown below

$$gSSR = \frac{\sum_{i_{total}=1}^{N_{total}} \left(\phi_e^{t-1}(i_{total}) - \phi_e^{t}(i_{total}) \right)^2}{\sum_{i_{total}=1}^{N_{total}} \left(\phi_e^{t-1}(i_{total}) \right)^2} \quad (1)$$

where extracellular potential Φ_e of element i_{total} in the global volume at time t and $(t-1)$, respectively and N_{total} is the total number of elements in the whole simulation.

We were also interested in the behavior of two further convergence criteria that could potentially reduce communication in the GS iterations. A local SSR (lSSR) disregarding the SSR values of other subvolumes would allow to compute the convergence criterion locally only and would not introduce a communication overhead for its computation. The lSSR is computed with

$$lSSR = \frac{\sum_{i_{sub}=1}^{N_{subvolume}} \left(\phi_e^{t-1}(i_{sub}) - \phi_e^{t}(i_{sub})\right)^2}{\sum_{i_{sub}=1}^{N_{subvolume}} \left(\phi_e^{t-1}(i_{sub})\right)^2} \quad (2)$$

where $N_{subvolume}$ is the total number of elements in the local subvolume and i_{sub} denotes the i^{th} element in the local subvolume.

As an intermediate case, we compute the global sum of local SSRs (gSlSSR) by

$$gSlSSR = \sum_{p=1}^{N_{proc}} \left\{ \frac{\sum_{i_{sub}=1}^{N_{subvolume}} \left(\phi_e^{t-1}(i_{sub}) - \phi_e^{t}(i_{sub})\right)^2}{\sum_{i_{sub}=1}^{N_{subvolume}} \left(\phi_e^{t-1}(i_{sub})\right)^2} \right\} \quad (3)$$

with the total number of processors N_{proc} and the total number of elements in the local subvolume $N_{subvolume}$. i_{sub} denotes the i^{th} element in the local subvolume. Thus the inter-processor communication during iteration can be reduced by factor 2 for the computation of the convergence criterion.

Figure 1 shows a flow chart of the computation and communication cycle of the computation. After each GS iteration, the ghost values of the extracellular potential for each subvolume need to be updated since Blue Gene is a distributed memory parallel supercomputer (see fig. 1). A further reduction in communication overhead can be achieved if this communication phase is not carried out at each iteration but for every fifth, tenth iteration or not at all. A calculation error might be introduced. We investigate this error by calculating a root mean square error (RMSE) of transmembrane voltage and extracellular potentials. The reference values of the extracellular potentials are given by the simulation with a communication phase after each GS iteration. At the end of each simulation, the extracellular potential is saved and the RMSE is computed.

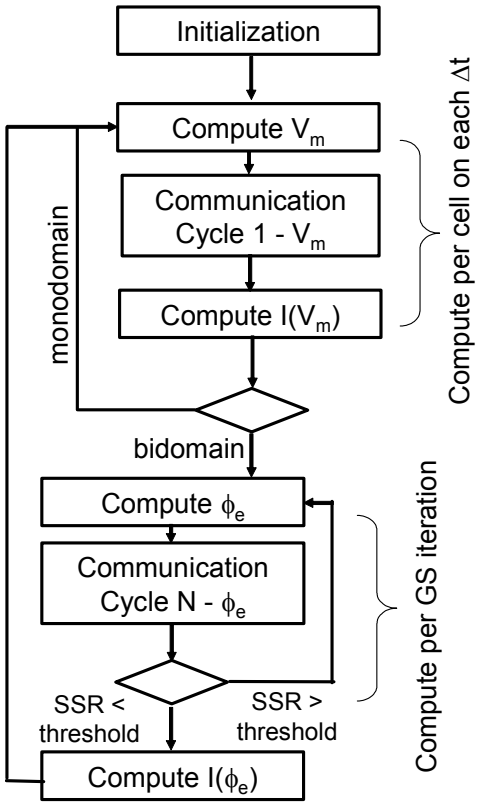

Fig. 1. Communication cycles in the simulations per time step. The transmembrane potential V_m is computed after initialization. The values of V_m at the boundaries between compute nodes are then communicated to compute the current density $I(V_m)$. Given the monodomain case, the V_m at the next time step is calculated. Given the bidomain case, the extracellular potential Φ_e is calculated by way of Gauss-Seidel method which requires an unknown number N of communication cycles depending on the number of iterations needed. This number is determined by either a maximum number of iterations or by checking the sum of squared residuals (SSR) against a threshold for convergence. Then the current density $I(\Phi_e)$ can be computed before starting the iteration for the next time step.

In simulation run of 1 sec requires 10000 steps fixed at time interval of 100 μs. The small time step is required for numerical stability of the cell model. The extracellular potential is calculated at each every fifth time step (a commonly used heuristic to reduce communication overhead).

The communication framework is build with standard MPI functions. The communication is synchronized by non-blocking send and receive functions used in

conjunction with the wait function. A communication list is created based on an orthogonal recursive bisection (ORB) algorithm that decomposes the tissues distributes to multiple processor in a manner to balance computational loads. The complete description of the data decomposition and communication is found in [5, 6]. For the computation of the SSRs, all-reduce functions were used since the gSSR and gSlSSR are functions over the whole data volume and not restricted to the subvolume allocated to each computational node.

3. Results

Figure 2 and 3 show the total run time as well as the respective computation and communication times for each simulation. If Φ_e is calculated every time step, the run times increase when reducing the communication phases during the GS iteration (fig. 2). A reduction of communication phases by factor 10 leads to an increase in run time by about factor 5.

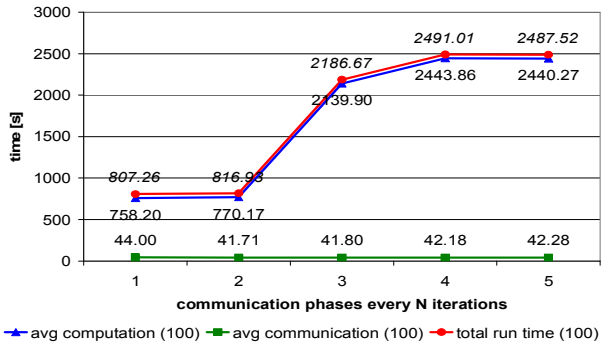

Fig. 2. Average computation and communication times and total run time as function of the number communication phases per GS iteration. The extracellular potential was computed every time step Simulations limited communication phase in the GS method each iteration, every 5th, 10th, 50th iteration or no communication. The maximum iteration number allowed is 100. The times are labels on the graph.

In contrast, the simulations where Φ_e was computed only every 5th time step show the expected result. A decrease in run time from 481 s to 415 s when the communication phases in the GS iteration was reduced (fig.3). The run time is also substantially reduced when the extracellular potential is computed every 5th versus each time step. This reduction can be seen by comparing total run times (red traces) in Figs. 2 and 3.

Counting the total number of GS iterations for a simulation, i.e. taking the sum of all GS iterations at each time step, the simulations in fig. 2 show an increase from around 23000 (communication phases every and every 5th GS iteration) iterations to over 10^5 iterations for the cases with communication phases less than every 10th GS iteration. While the communication phase is reduced in the latter cases, the computation overhead leads to the increase in overall run time since the simulation is computation bound with little communication overhead. The simulations in fig. 3 show a decrease in total number of GS iterations from about 11000 to less than 8000.

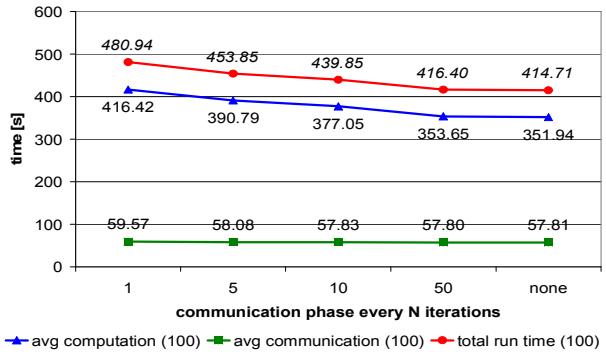

Fig. 3. Average computation and communication times and total run time as function of the number communication phases per GS iteration. The extracellular potential was computed every 5th time step Simulations limited communication phase in the GS method each iteration, every 5th, 10th, 50th iteration or no communication. The maximum iteration number allowed is 100. The times are labels on the graph.

Note that the results are essentially the same whether the maximum number GS iterations per time step is limited to 100 or 1000. This effect is investigated in Fig. 4. Here the number of GS iterations per time step is shown for the first 0.1 s in the simulation; the different color traces correspond to limiting inter-processor communication to every 5th, 10th, 50th GS iteration or no communication as labeled. The number of GS iterations quickly falls below 20 and seldom exceeds 10 after 20 ms simulation time as shown Fig. 4. The solution converges at every time step when limiting the maximum number of iterations to 1000.

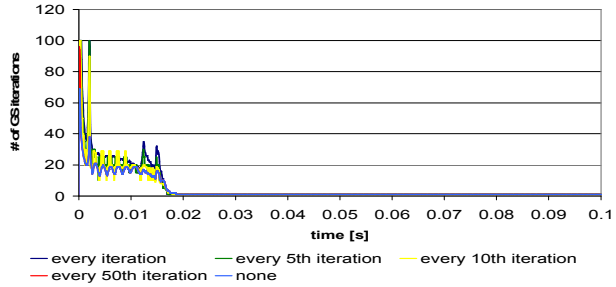

Fig. 4. Iteration count per time step with maximum 1000 iterations allowed for the simulation that the extracellular potential is computed every time step.

The lSSR has overall the same magnitude as the gSSR (fig. 5 and 6). The gSlSSR is magnitudes larger than both lSSR_ind and gSSR. Although the gSSR drops even lower than convergence threshold and remains below the lSSR, the latter also achieves values below the convergence threshold. However, it is interesting to note that while the gSSR remains low, the lSSR at times shows spikes and increases in values during the simulation.

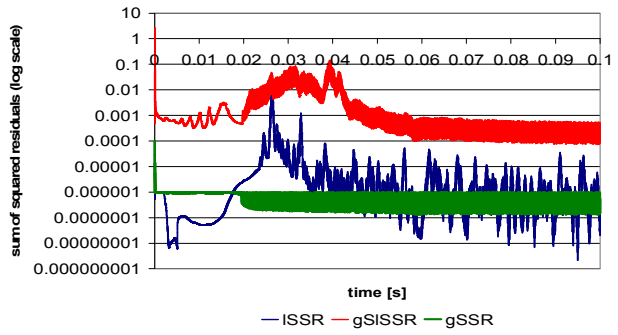

Fig. 5. This figure shows the lSSR (blue), gSlSSR (red) and gSSR (green) for the simulation with 100 iterations allowed and computation of Φ_e at every time step.

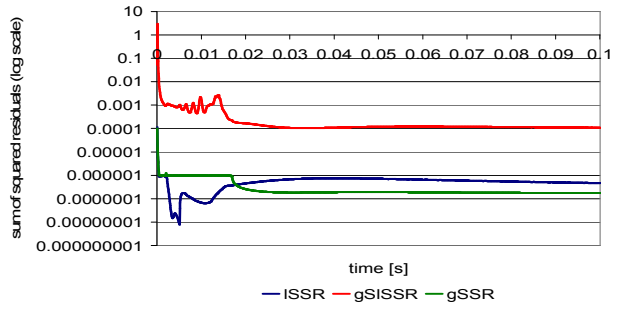

Fig. 6. This figure shows the lSSR (blue), gSlSSR (red) and gSSR (green) for the simulation with 100 iterations allowed and computation of Φ_e at every 5th time step.

The RMSE of the transmembrane voltage V_m and the extracellular potential Φ_e were increased when reducing communication phases in the GS iterations. While reducing them by factor 10 the RMSE of V_m is of the order of 10^{-5}. A reduction by factor 50 or more increases the error to 0.105. Increasing the maximum GS iteration steps did not change that result. The RMSE of Φ_e was in the order of 10^{-3} and a reduction in communication phases lead to an higher RMSE score in the order of 10^{-2}. Computing Φ_e every 5 times steps produces a better result when the number of communication phases is also reduced. In his case, the RMSE values are in the order of by 10^{-4} and 10^{-3} for V_m and Φ_e respectively.

Since communication is less expensive then computation with respect to wall clock time, an increase in communication has not a large effect on the overall run time.

4. Discussion and conclusions

While preliminary, our results illustrate that bidomain computation is feasible on a fully distributed memory system. Moreover, the results suggest some heuristic approaches may further reduce the inter-processor communication to improve the execution time of large-scale simulations. Especially the local convergence criterion seems promising. Since the global convergence criterion usually remains below the local criterion after the initial phases of the simulation, the GS method can be assumed to converge if the local convergence is guaranteed. Moreover, the results suggest some heuristic approaches may further reduce the inter-processor communication to improve the execution time of large-scale simulations.

Future work will focus on simulating large anatomical data-sets on system sizes greater than 1024 processors to enable fast simulation of cardiac models on organ level including detailed electrophysiological and models.

References

[1] Vigmond EJ, Weber dos Santos R, Prassl AJ, Deo M, Plank G. Solvers for the cardiac bidomain equations. Prog Biophys Mol Biol. 2008;96(1-3):3-18

[2] Ten Tusscher KHWJ, Noble D, Noble PJ, Panfilov AV. A model for human ventricular tissue. Am J Physiol. 2004;286:H1573-H1598.

[3] Saleheen HI, Ng KT. A New Three-Dimensional Finite-Difference Bidomain Formulation for Inhomogeneous Anisotropic Cardiac Tissues. IEEE Trans Biomed Eng. 1998;45(1):15-25

[4] Potse M, Dubé B, Richer J, Vinet A, Gulrajani RM. A Comparison of Monodomain and Bidomain Reaction-Diffusion Models for Action Potential Propagation in the Human Heart. IEEE Trans Biomed Eng. 2006;53(12):2425-2435

[5] Reumann M, Fitch BG, Rayshubskiy A, Keller DUJ, Weiss DL, Seemann G, Dössel O, Pitman MC, Rice JJ. Large Scale Cardiac Modeling on the Blue Gene Supercomputer. Conf Proc IEEE Eng Med Biol Soc. 2008;2008:577-580

[6] Nyland L, Prins J, Yun RH, Hermans J, Kum H-C, Wang L. Achieving scalable parallel molecular dynamics using dynamic spatial decomposition techniques. Journal of Parallel and Distributed Computing, 1997;47(2): 125-138

Address for correspondence

Name Matthias Reumann
Full postal address IBM T. J. Watson Research Center, 1101 Kitchawan Road, Yorktown Heights, NY, 10598, USA
E-mail address mreumann@ieee.org