HIGH-FIDELITY INJECTOR MODELING WITH PARALLEL FINITE ELEMENT 3D ELECTROMAGNETIC PIC CODE PIC3P *

A. Candel†, A. Kabel, L. Lee, Z. Li, C. Ng, G. Schussman and K. Ko,
SLAC, Menlo Park, CA 94025, U.S.A.

Abstract

SLAC's Advanced Computations Department (ACD) has developed the parallel Finite Element 3D electromagnetic code suite ACE3P for modeling of complex accelerator structures. The Particle-In-Cell module Pic3P was designed for simulations of beam-cavity interactions dominated by space charge effects. Pic3P solves the complete set of Maxwell-Lorentz equations self-consistently and includes space-charge, retardation and boundary effects from first principles. In addition to using conformal, unstructured meshes in combination with higher-order Finite Element methods, Pic3P also uses causal moving window techniques and dynamic load balancing for highly efficient use of computational resources. Operating on workstations and on leadership-class supercomputing facilities, Pic3P allows large-scale modeling of photoinjectors with unprecedented accuracy, aiding the design and operation of next-generation accelerator facilities. Applications include the LCLS RF gun.

THE PARALLEL CODE PIC3P

In Pic3P, the full set of Maxwell's equations is solved numerically in time domain using parallel higher-order Finite Element methods. Electron macro-particles are pushed self-consistently in space charge, wake- and external drive fields.

Finite Element Time-Domain Field Solver

Ampère's and Faraday's laws are combined and integrated over time to yield the inhomogeneous vector wave equation for the time integral of the electric field **E**:

$$\left(\varepsilon\frac{\partial^2}{\partial t^2}+\sigma\frac{\partial}{\partial t}+\nabla\times\mu^{-1}\nabla\times\right)\int^t\mathbf{E}(\mathbf{x},\tau)\,\mathrm{d}\tau=-\mathbf{J}(\mathbf{x},t), \tag{1}$$

with permittivity ε and permeability μ. The effective conductivity σ provides a simple model for Ohmic losses.

The computational domain is discretized into curved tetrahedral elements and $\int^t\mathbf{E}\,\mathrm{d}\tau$ in Equation (1) is expanded into a set of hierarchical Whitney vector basis functions $\mathbf{N}_i(\mathbf{x})$ up to order p within each element:

$$\int^t\mathbf{E}(\mathbf{x},\tau)\,\mathrm{d}\tau=\sum_{i=1}^{N_p}e_i(t)\cdot\mathbf{N}_i(\mathbf{x}). \tag{2}$$

For typical simulation runs with second-order elements (curved and using second-order basis functions), $N_2 = 20$. Up to $N_6 = 216$ different basis functions can be used in each element. Tangential continuity between neighboring elements reduces the global number of degrees of freedom, in contrast to discontinuous Galerkin methods.

Substituting Equation (2) into Equation (1), multiplying by a test function and integrating over the computational domain results in a system of linear equations (second-order in time) for the coefficients e_i. Numerical integration is performed with the unconditionally stable implicit Newmark-Beta scheme [1]. More detailed information about the employed methods has been published earlier [2].

Higher-Order Particle-Field Coupling

Electron macro particles are specified by position **x**, momentum **p**, rest mass m and charge q. The total current density **J** in Equation (1) is then approximated as

$$\mathbf{J}(\mathbf{x},t)=\sum_i q_i\cdot\delta(\mathbf{x}-\mathbf{x}_i(t))\cdot\mathbf{v}_i(t), \tag{3}$$

for delta-particles with $\mathbf{v}=\frac{\mathbf{p}}{\gamma m}$, $\gamma^2=1+|\frac{\mathbf{p}}{mc}|^2$. The classical relativistic collision-less Newton-Lorentz equations of motion are integrated using the standard Boris pusher [3].

Starting with consistent initial conditions and fulfilling the discrete versions of Equation (1) and the continuity equation

$$\frac{\partial\rho}{\partial t}+\nabla\cdot\mathbf{J}=\mathbf{0} \tag{4}$$

simultaneously during time integration leads to numerical charge conservation.

The use of higher-order Finite Elements not only significantly improves field accuracy and dispersive properties [4], but also leads to intrinsic higher-order accurate particle-field coupling. For delta-particles, the numerical current deposition involves the (exact) evaluation of line integrals over vector basis functions along the elemental particle trajectory segments for a given time step, using Gaussian quadrature [5].

* Work supported by the U. S. DOE ASCR, BES, and HEP Divisions under contract No. DE-AC002-76SF00515.
† candel@slac.stanford.edu

Code Performance Optimization

Pic3P has been optimized for high efficiency both on workstations as well as leadership-class supercomputers. In the following, two specific methods to optimize the parallel performance of Pic3P are introduced, i.e. the causal moving window technique and dynamic parallel load balancing.

Causal Moving Window Field computations can be restricted to the causality region around the particle bunch without any loss of accuracy, assuming the (approximate) trajectory of the particles through the gun is known. The "moving window" feature of Pic3P makes use of adaptive p-refinement where the Finite Element basis function order can be specified for every element, including the omission of certain elements. This can save orders of magnitude in memory and CPU time requirements and allows full EM 3D PIC simulations on desktop computers [6].

Dynamic Load Balancing Pic3P is designed to scale to large problem sizes, and both the fields and the particles are partitioned using a specialized load balancing scheme. As in ACE3P's time-domain code T3P, the computational Finite Element mesh and the corresponding field degrees of freedom are typically partitioned using graph-based methods (e.g., with ParMETIS) [7]. The macro-particles are typically partitioned using a geometric "recursive coordinate bisection" method (e.g., with Zoltan). Figure 1 shows a typical example of the field and particle partitioning. The absence of explicit locality of particle and field data

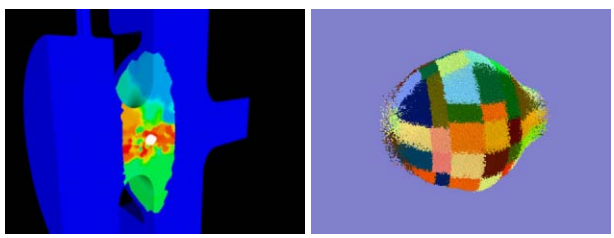

Figure 1: Parallel dynamic load balancing in Pic3P for a typical LCLS RF gun PIC simulation. Partitioning of fields (left) and particles (right) onto different processors is indicated by colors. On the left, the particle bunch is shown in white while the omitted field region from the causal moving window scheme is shown in blue.

necessitates a sophisticated parallel particle-field coupling scheme, which is outlined in the following. First, the geometric domain (such as a collection of bounding boxes) containing all particles is collectively obtained by all processors. Second, every process determines whether it owns or needs fields for the particle region for the next particle pushing step. Then, MPI communicators are constructed accordingly, connecting sub-groups of involved processes, and the "PIC Mesh", which consists of tetrahedral elements and corresponding field coefficients, is collected with collective MPI operations such as `Allgatherv`. Finally, the particles are pushed using the (now locally available) PIC Mesh information. The current deposition "scatter" step is done similarly.

By segmenting the particle region into fewer or more pieces, the memory and time requirements during the communication steps can be traded, such that both strong and weak optimal scalability can be achieved. The execution order of PIC Mesh updates for different communicators is optimized and re-ordered such that disjoint sub-groups of processes can communicate simultaneously. By slightly enlarging the bounding box around the particles (by extrapolating the trajectories), a new PIC Mesh only needs to be created from time to time (e.g., every 50 time steps), and re-using the same objects and communicators leads to significant savings in computational resources, as only the changing degrees of freedom for electromagnetic fields and particle currents need to be communicated.

This dynamic load balancing scheme allows the solution of large problems with hundreds of millions of field degrees of freedom (DOFs) and billions of particles and enables unprecedented accuracy in self-consistent state-of-the-art simulations of beam-cavity interactions on modern supercomputers. Table 1 shows typical runtime parameters for LCLS RF gun PIC simulations performed with Pic3P on the Jaguar Cray XT5 machine at NCCS. For small problem sizes (i.e., up to a few million DOFs, and less than 100 CPUs), the use of direct factorization methods for the field solution is most efficient. For larger problems, iterative solvers are used, since they display much better scalability. For typical problem sizes, the performance of the memory subsystem is most important. Due to the small relative size of the particle bunch compared to the overall field domain in this LCLS RF gun application, the use of more than a few thousand CPUs is not efficient in the current implementation, as the global PIC Mesh construction and communication begins to strain the network subsystem. However, the current scheme could be easily modified to only use a subset of the processors for particle pushing. This should lead to better scalability for extremely large problems with highly localized particle distributions, similar to the good scalability experienced for small to mid-size problems.

Table 1: Typical runtime parameters on a Cray XT5 for LCLS RF gun emittance calculations with Pic3P. For the first case, causal moving windowing and a direct solver were used. For the other cases, full-domain calculations were performed with an iterative field solver (i.e. , CG with incomplete Cholesky preconditioner).

CPUs	Particles	Elements	DOFs	Walltime/step
12	100k	305k	53k	0.23 secs
120	1M	305k	2M	1.3 secs
1200	10M	2.8M	17.7M	2.4 secs
12,000	100M	23M	142M	24 secs

VALIDATION: LCLS RF GUN

Benchmark PIC simulations of the 1.6-cell S-band LCLS RF gun are presented in the following [8]. Simulation parameters are: π-mode, 120 MV/m, 1 nC, 10 ps, 1 mm beer-can initial bunch distribution, centroid injection phase -58° and no solenoid. These parameters allow comparisons between the 3D results from simulations with Pic3P and PARMELA and the 2D results from simulations with Pic2P and MAFIA.

Figure 2 shows a comparison of transverse emittance results by the different codes. For Pic3P simulations, a conformal, unstructured 3D (1/4) mesh model with only 305k tetrahedral elements is sufficient to reach convergence, with mesh refinement along the center of the beam pipe. High fidelity cavity mode fields are obtained with the parallel Finite Element frequency domain code Omega3P and directly loaded into Pic3P as drive fields.

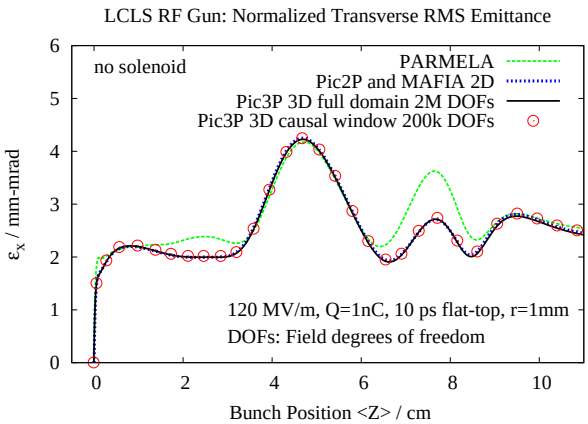

Figure 2: Comparison of normalized transverse RMS emittance as a function of beam position in the LCLS RF gun as calculated with PARMELA, Pic2P and MAFIA 2D (both agree), and Pic3P, where the causal moving window technique reduces the problem size by one order of magnitude.

Excellent agreement between 3D results from Pic3P and the 2D results from Pic2P is found, as expected from the high cylindrical symmetry in the fields, as well as perfect agreement with MAFIA 2D as expected from the convergence behavior of the codes. PARMELA results differ as space-charge effects are significant, presumably because wakefield and retardation effects are ignored, as detailed in a previous study [2]. Simulation results starting from a measured initial bunch distribution have been published earlier [6].

SUMMARY

SLAC has developed the first parallel higher-order Finite Element 3D PIC code Pic3P, for realistic modeling of space-charge dominated beam-cavity interactions. As a part of the ACE3P code suite, Pic3P uses state-of-the-art parallel Finite Element methods on conformal, unstructured meshes. Pic3P has been optimized for large-scale modeling on modern supercomputers and causal moving window techniques and dynamic load balancing enable high-fidelity simulations of electron injectors with unprecedented efficiency and accuracy. Applications include emittance calculations for the LCLS RF gun. Benchmark simulations show excellent agreement between Pic3P and MAFIA TS2 results, while results obtained with the electrostatic code PARMELA show some differences, presumably because wakefields and retardation effects are omitted.

ACKNOWLEDGMENTS

This work was supported by the US DOE ASCR, BES, and HEP Divisions under contract No. DE-AC02-76SF00515. This research used resources of the National Energy Research Scientific Computing Center, and of the National Center for Computational Sciences at Oak Ridge National Laboratory, which are supported by the Office of Science of the U. S. Department of Energy under Contract No. DE-AC02-05CH11231 and No. DE-AC05-00OR22725. – We also acknowledge the contributions from our SciDAC collaborators in numerous areas of computational science.

REFERENCES

[1] N. M. Newmark, “A method of computation for structural dynamics”, Journal of Eng. Mech. Div., ASCE, vol. 85, pp. 67-94, July 1959.

[2] A. Candel et al., “Parallel Higher-order Finite Element Method for Accurate Field Computations in Wakefield and PIC Simulations”, Proc. ICAP 2006, Chamonix Mont-Blanc, France, Oct. 2-6, 2006.

[3] J. P. Boris, “Relativistic plasma simulation-optimization of a hybrid code”, Proc. Fourth Conf. Num. Sim. Plasmas, Naval Res. Lab, Wash. D.C., pp. 3-67, Nov. 2-3, 1970.

[4] M. Ainsworth, “Dispersive properties of high-order Nedelec/edge element approximation of the time-harmonic Maxwell equations”, Philos. trans.-Royal Soc., Math. phys. eng. sci., vol. 362, no. 1816, pp. 471-492, 2004.

[5] J. W. Eastwood, W. Arter, N. J. Brealey, R. W. Hockney, “Body-fitted electromagnetic PIC software for use on parallel computers”, Computer Physics Communications, vol. 87, Issues 1-2, Particle Simulation Methods, pp. 155-178, May 2, 1995.

[6] A. Candel et al, “High-Fidelity RF Gun Simulations with the Parallel 3D Finite Element Particle-In-Cell Code Pic3P”, AIP CP1149, Proc. Workshop on Polarized Electron Sources, pp. 1114-1118, Oct 1-3, 2008

[7] A. Candel et al., “State of the art in electromagnetic modeling for the Compact Linear Collider”, Proc. J. Phys.: Conf. Ser., vol. 180, pp. 012004 (10pp), 2009.

[8] L. Xiao et al., “Dual Feed RF Gun Design for the LCLS”, Proc. PAC 2005, Knoxville, Tennessee, May 15-20, 2005.