

Paper for 3BEAMS May 98 Chicago

Rigorous Simulation of Statistical Electron-Electron Interactions with Fast Multipole Acceleration and a Network of Workstations

Victor S.H. Wen, Owen T. Carmichael*, Hiroshi Yamashita** and Andrew R. Neureuther

Electronics Research Laboratory, Department of Electrical Engineering and Computer Sciences, University of California, Berkeley CA 94720-1771, Tel:(510) 642-4590 Fax:(510) 642-2739, neureuth@eecs.berkeley.edu

* CS Graduate Student at CMU

** ULSI Device Development Laboratories, NEC Corporation, 1120 Shimokuzawa, Sagamihara, Kanagawa 229, Japan

ABSTRACT

Rigorous simulation, which in concept is nearly linear in CPU time with micro-Ampere beam currents, has been developed for analyzing statistical electron-electron interactions in lens free regions joined by ideal lenses and apertures. The simulation method is based on combining fast multipole acceleration (FMA) with modifications to a message passing interface for use on multi-processor and network of workstations (NOW). As test cases, the distribution of scattered electrons versus deviation from focus is given as well as results for the interaction between simultaneously imaged spots. The performance improvement per level of spatial division was only about a factor of 2 instead of the expected factor of 8 due to the pencil nature of the beam. An FMA method that adapts to beam density and geometry is needed. It is projected to be possible to simulate a 30uA beam current on a quad-Pentium Pro machine in half a day and in matter of hour on a network of workstation.

I. Introduction

Projection electron-beam systems such as cell projection and SCALPEL are candidates for the next generation lithography tools [1,2]. Beam blur due to interactions which cannot be compensated by a focus correction create a fundamental throughput versus beam current limit in electron-beam lithography. Absorption or modifications of the trajectories of the electrons in the masked region are used to create contrast between exposed and unexposed regions on the wafer. Of special concern are additional effects of scattered electrons from mask edges in stencil masks [3] and the energy and angle broadening from substrate scattered electrons in SCALPEL. These effects are of greatest concern for electrons having a combination of position and velocity to follow nearly similar trajectories through the electron optical system and as a consequence perturb the main body of electrons that form the ideal image.

To date only approximate statistical theories are able to help understand these limits in statistical electron-electron interactions. For conventional electron-beam columns modeling techniques based on coherence length of sets of electrons have been introduced by Jansen [4]. The analysis has been generalized to the SCALPEL system by Mkrtchyan [5]. While the assumption of coherence length makes the analysis tractable, this fundamental assumption has yet to be investigated by direct methods. Codes for simulation of electron beam optics are widely available and have matured to the point of commercialization [6]. Codes typically require several hours to simulate effects of currents up to 1uA. If brute force methods were to be added to examine interactions between every pair of electrons the simulation of beam currents up to 30uA might be

expected to take some 900 times longer. Fortunately, Greengard introduced computational methods for gravitational fields [7] based on fast multipole acceleration (FMA) in which simulation time increases nearly linear with the number of particles. Codes with FMA analysis are being developed such as the DPMTA package available from Duke University [8]. Parallelization is also possible to speed up the simulation and can be implemented by means of special interfaces such as the message passing interface (MPI) available from Mississippi State University for PC-NT's [9] and for the Network of Workstations (NOW) at Berkeley [10].

This paper describes results from parallelizing and applying the DPMTA code to simulate statistical electron-electron interactions in lens free regions joined by ideal lenses and apertures. Section II describes the simulation region geometry, the codes and the master/slave organizational concepts. Section III gives several physical examples used in testing, including spot size growth with thermal energy and interactions between crossing beams. Data on the performance of the FMA algorithm for very large numbers of electrons is described in Section IV.

II. Method

Geometry

The basic setup of the electron-beam simulator consists of a cylindrical tube with a focusing lens on one end and a collector plate on the other as shown in Figure 1. At every iteration, new electrons, the number of which depends on current density, are emitted from the lens. Forces between existing electrons in the simulation region are also

calculated and the electrons are moved according to the total force they experience. In our simulation studies, a domain length of 20cm was used corresponding to the lens free region which might be used in a electron projection system such as the Hitachi Cell Projection System. A domain radius of 10um, giving it a very high aspect ratio, was used performance benchmarking. In the beam blur studies, a radius of 20cm was used to make more effective use of the FMA.

Sources

The electron beam simulator has a set of internal source models as well as accepts electron injection data from a file. Two internal models are a single point source and twin point sources of electrons. Both models have source temperature as a user-adjustable parameter. The point source model with temperature is used to investigate electrons with “highway-like” collisions due to thermally excited electrons hitting slower-travelling ones on similar trajectories. The twin point source is capable of simulating the formation of two simultaneous images, each with different current density, focus locations and accelerating voltages. The purpose is to investigate the extent of beam blur due to inter-beam effects.

The simulator also accepts electron mask data and employs a projection source that starts at the mask level. Scattered electron mask data from Monte Carlo simulation is read in from disk[11]. Electrons are generated according to the exit position, angle, energy from the data. Transmitted electrons through the mask can also be simulated. The electrons are then focused by an ideal lens, passed through a limiting aperture, and finally absorbed by

the target plane. The projection electron simulator is useful in investigating the perturbation induced by scattered electrons from mask and mask edges on the ideal image.

Code

The FMA algorithm divides the simulation region into smaller, equal-sized sub-regions, the number of which depends on the spatial decomposition level. For example, with level 1 spatial decomposition, the number of sub-regions created is 8, with level 2 it is 64. In general, the number of sub-regions created is equal to 8^n , where n is the decomposition level. The pair-wise force calculations between electrons occur only in the same sub-region. For all sub-regions, a multipole moment is calculated, based on the spatial configuration of electrons in a given sub-region. These multipole moments are passed to all other sub-regions. The algorithm then uses the multipole moments to adjust the total force experienced by each electron.

The FMA package we utilized, called DPMTA (Distributed Parallel Multipole Tree Algorithm), in our electron beam simulators, is written by the group of William T. Rankin of Duke University. The package uses Parallel Virtual Machine (PVM) as its distributed framework. Modifications were made to use the Message Passing Interface protocol to facilitate better parallelization on the Berkeley NOW cluster. Because of the portability of MPI, the same code for the NOW cluster can also run on shared-memory multiprocessor PC running NT with MPI/NT library from Mississippi State University.

Master-Slave Setup

The simulator employs a master-slave setup (Figure 2), resulting in the total number of processes needed to be 2^n+1 . Each slave process receives the electrons in its region of simulation and performs the task of direct calculation of forces between electrons, passing multipole moments to neighboring slaves if needed, and sending the resultant force vectors back to the master process. The master process is responsible for allocating space for electron positions and force vectors, applying forces to electrons, generating new electrons and outputting electron positions. The communication between master and slaves involves sending and receiving electron position and force data. This time goes up linearly with the number of electrons. The communication between slaves involves passing multipole moments to and from different boxes, which grows exponentially with higher levels of spatial decomposition.

Idealized lenses and apertures can be included at reference planes where electron trajectories are modified. A web interface is being developed for specifying simulation parameters and viewing results. A format is being developed for injecting in and capturing from the simulation domain, such that this simulator can be invoked through the web by electron optical design programs as an additional element in the column. It will be accessible through the Berkeley LAVA web site at <http://cuervo.eecs.berkeley.edu/Volcano>.

III. Physical Examples

Basic Testing

The simulator was tested initially by confirming that the beam focuses on a single spot when the electrostatic force is turned off and temperature is set to zero. Then the force was enabled and pairs of electrons were emitted simultaneously and the separation distance at the target plane was checked by thorough analysis. Next the blooming at the beam spot was checked. Figure 3a is for a case of a 3uA beam current starting from a radius of 20cm and focus to an infinitesimal spot 20cm away. Beam blooming takes place due to the interactions between electrons simultaneously present in the column. It is relatively small due to the large separation between electrons except near the focus spot. The diameter is, judged by inspection, to be about a 0.3um. Figure 3b is for a 1uA current starting from a radius of 10um and focussed to an infinitesimal spot 20cm away. In this case, beam blooming is much larger due to small spacing between electrons throughout the column. It is, judged by inspection, to be about 20um in diameter. This is in good agreement with a simulation of the Boersch effect of 21um (3 sigma) as calculated by the electron beam software of Munro[6].

Longitudinal Velocity

An interesting case to illustrate statistical interactions is a hypothetical one of electrons having different thermal velocities progressing down the simulation region. The exit direction vector of each electron from the lens is tuned so that, without interaction with other electrons, all would pass through the focus point. A source temperature can be used to control the relative variation in velocities. This temperature is then adjusted so that an

expected number of "highway-like" collisions by overtaking electrons occur. At 1uA, for example, a temperature of 2000 K causes about 20 collisions to occur. Figure 4 shows results for such a simulation with a beam convergence angle of 0.005mrad. With the temperature on, the number of electrons scattered outside the 2um radius increases from 56.7% to 62.5% as shown in Figure 5a and 5b. Studies of density versus radius and distributions in energy and direction can give insight into scattering mechanisms and can be used to test models which aggregate electrons in to groups which travel intact for a coherence length.

Crossing Beam

The ability of the eBeam simulator to simulate two point sources simultaneously facilitates the study of the perturbation of one spot due to the presence of additional current forming a separate pattern. An example is shown in Figure 5 for two 1uA spots formed by an ideal lens with radius equals to 20cm. About half of the scattered electrons occur in pairs indicating that only half of the scattering is due to the presence of the other beam. In Figure 6 the larger spot has a beam current of 10uA. The small spot has now increased in size from about 10 nm to 50 nm by the presence of the second beam. Hardly any pairs of scattered electrons are noticeable. The simulator is thus well suited for analyzing scattering among electrons forming different elements in a projected pattern.

IV. Performance

The speed up expected from FMA algorithm was not fully realized when the number of particles was increased. In Figure 7, the run times for 100 time steps for 30000 particles

with levels of spatial decomposition from two to seven is presented. The pair-wise force calculation time (the “Direct” row) decreases by factor of two with each increment of decomposition level. However, the number of particles per box supposedly decreases by a factor of eight for the same transition. The data shows the reduction in time as if the number of electrons per box decreases by factor of two only! This phenomenon is due to the high concentration of electrons near the column axis (pencil shape) when it is focused. This region is only divided in two for each level, leaving many electrons in some boxes requiring pair-wise force calculations.

To test this focused pencil beam hypothesis, simulation region with focus at infinity was introduced. As Figure 9 demonstrates, the run times decrease substantially. For 300000 particles (30uA), the total run time per slave is 9680 seconds for 10 iterations at level 6 (96800s for 100 iteration), as compared with 33000 seconds for 30uA with simulation conditions similar to that from Figure 8. Moreover, we now are able to observe an almost eight fold decrease in direct calculation time as the decomposition level increases. For example, the pair-wise force calculation time decreases from 193s to 34s as the decomposition level decreases from four to five for the 30000 particles (3uA) case.

The two dominant overhead operations involved in employing DPMTA occur in `Recv_Particles()` and `MultipoleResize()` functions. These two routines grow linearly with the number of particles and also with higher levels of decomposition. Thus, as the time for pair-wise force calculation decreases, the time for overhead increases. The point, at

which these two curves cross, will be the sweet spot, the optimum level of decomposition for the given physical setup.

From the above test data, an empirical approach has also been developed to model and optimize the total calculation time per iteration as a function of the number of electrons (N_e) and level of decomposition (N_L).

The direct calculation time is modeled by:

$$T_d = k_d N_e^2 * 8^{-N_L} \quad (1)$$

$$(k_d = 3.6 * 10^{-4} \text{ sec}/N_e^2)$$

The overhead run time per iteration is modeled by:

$$T_o = k_o N_L * 8^{N_L} \quad (2)$$

$$(k_o = 2.1 * 10^{-8} \text{ sec}/N_e)$$

The total run time, T_{tot} , is the sum of T_d and T_o . The optimal decomposition level can be found by taking derivative of T_{tot} with respect to N_L and setting it equal to zero. For the simulations given above, the optimum number of electrons per box ($N_{e,opt}$) is 4.2. In general, the optimum level of decomposition and total run time is given by equations 3 and 4.

$$N_{L,opt} = \frac{1}{2 \ln 8} \ln \left(\frac{k_d N_L}{k_o} \right) \quad (3)$$

$$T_{opt} = 2 \sqrt{k_d k_o N_L^3} \quad (4)$$

As a check, for 300000 particles or 30uA, $N_{L,opt} = 5.3$ and $T_{opt} = 900 \text{ sec/iteration}$ which agrees closely with run time from Figure 8.

Three important conclusions can be drawn from our data:

1. The low optimal average number of electrons per box (4.2) means that both the FMA computation and message passing are very efficient. The message passing may be particularly efficient due to the shared-memory of the quad-Pentium Pro on which the data was taken.
2. The projected run time for 300000 electrons or 30uA will be 900sec per iteration. For a typical run of 260 iterations, only half a day is needed on the quad-Pentium Pro. However, it will require development of FMA that efficiently handles the high aspect ratio of the electron optical column shape and the wide variation in charge density at crossovers.
3. In limited testing with existing NOW, with a slower processor and up to 8 machines, the results were similar to a multi-processor machine. This extrapolates to simulation time of about 30 minutes for 30uA on 33 machines.

Conclusion

Fast multipole acceleration (FMA) and parallelization have been used to develop a simulator for statistical electron-electron interactions in lens free regions joined by ideal lenses and apertures. The distribution of scattered electrons versus deviation from focus is given as well as results for the interaction between simultaneously imaged spots. The program is being made available on the web for testing assumptions used in classical analysis methods. The performance gain for a tight and focused electron column with higher spatial division level was only a factor of 2 instead 8 due to the lack of uniform distribution of electrons among cells with each subdivisions. Testing on more cubic and uniformly filled cells indicates that in the future a locally adaptable FMA could be used

to simulate a system with a 30uA beam current in half a day on a quad-Pentium Pro computer or 30 minutes on NOW.

References

- [1] Y. Nakayama, et al, J. Vac. Sci. Technol. B **8** (6), 1836 (1990)
- [2] S. D. Berger, et al, J. Vac. Sci. Technol. B **9** (6), 2996 (1991)
- [3] H. Yamashita, E. Nomura and H. Nozue, *Effects of Accelerating Voltage and Pattern Size on Electron Scattering by EB Mask*, J. Vac. Sci. Technol. B **15** (6), (1997)
- [4] G. H. Jansen, J. Vac. Sci. Technol. B **6**, 1977 (1988)
- [5] M. M. Mkrtchyan, J. A. Liddle, S. D. Berger, L. R. Harriott, A. M. Schwartz and J. M. Gibson, *Stochastic scattering in charged particle projection systems: A nearest neighbor approach*, J. Appl. Phys. **78** (12)
- [6] E. Munro, Nucl. Instrum. Methods Res. A **258**, 443 (1987)
- [7] L. Greengard. *The Rapid Evaluation of Potential Fields in Particle Systems*. MIT Press, Cambridge, MA, 1988
- [8] Parallel Multipole Tree Algorithms from the Scientific Computing Group Duke University <http://www.ee.duke.edu/Research/SciComp/software.html>
- [9] MPI for PC-NT provided by The National Science Foundation Engineering Research Center for Computational Field Simulation at Mississippi State University <http://www.erc.msstate.edu/>
- [10] The Berkeley Network of Workstations (NOW) project <http://now.cs.berkeley.edu/>