SPIN TRACKING WITH GPUS TO 250 GeV IN RHIC LATTICE

V.H. Ranjbar, Tech-X Corp, Boulder, CO 80303
F. Meot, M. Bai, BNL, Upton, NY 11973

Abstract

We have benchmarked UAL-SPINK against Zgoubi and a list of well understood spin physics results. Along the way we addressed issues relating to longitudinal dynamics and orbit bump and distortion handling as well as appropriate slicing necessary for the TEAPOT-SPINK spin orbit integrator. We have also ported this TEAPOT-SPINK algorithm to the GPU's. We present the challenges associated with this work.

TEAPOT-SPINK INTEGRATOR

Teapot [1] uses a thin element kick plus drift formalism to integrate simplectically a particles trajectory through a lattice of magnetic elements. Accuracy is achieved by dividing up a given element into smaller and smaller drift-kick-drift units.

SPINK [3] generates spin transport by solving the T-BMT equation for a thin spin precessing magnetic element. It accomplishes this by calculating 1st order spin transport matrix for a thin element using the orbit phase space coordinates to determine the integrated value of the magnetic field across a given element. In the original version of SPINK the average of the initial and final phase space values across a thick element where used to derive the integrated field strength and derive the spin transport map. The 1st order orbit maps used were first calculated in MAD[4] and then imported to SPINK.

In this new code transport maps are generated using the phase space values at each spin kick location. No averaging is done. The usual Teapot orbit transport 'units' were further divided in half with a spin kick applied in the middle. This approach preserves both the symplecticity and unitarity of the orbit and spin respectively.

The current structure is held together in the Unified Accelerator Library (UAL) [6] framework using it to read in the lattice values and beam parameters.

BENCHMARKING

To establish confidence with the new code we decided on several benchmarking tests for which we had well understood physics results. These tests included effects on spin tune Q_s and crossing of intrinsic resonances and snake resonances. We also compared our results against more exact power series integrator code Zgoubi [5]. Table 1 shows a list of results for both TEAPOT-SPINK and Zgoubi.

In Fig. 1 the normal intrinsic resonance crossing is shown for two different intrinsic resonances in RHIC.

In Fig. 2 we show the odd and even snake resonance crossings. Odd should exist with even in the absence of orbit errors unlike even snake resonances which appear only in the presence of errors.

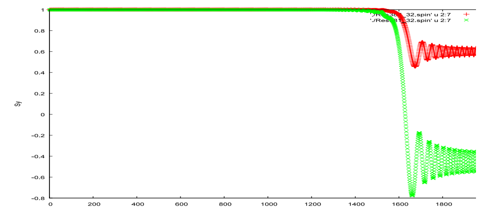

Figure 1: Sample intrinsic resonance crossing without snakes at $G\gamma = 381.32$ (green) and 387.32 . Depol [2] estimates resonance strength at 10 π mm-mrad 95 rms of 0.133 and 0.055 respectively, while TEAPOT-SPINK estimates 0.129 and 0.052 .

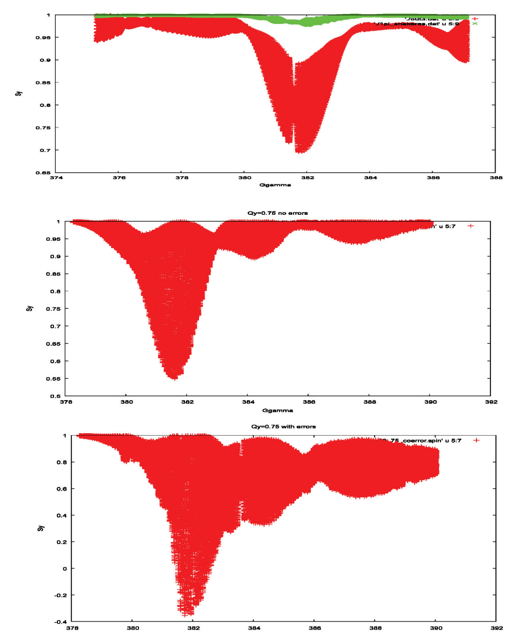

Figure 2: (Top plot: 7/10 odd order snake resonance averaged over 8 particles at .1 (green) and 1PI (red) with Qy = 0.7. Middle plot: even order snake resonance no errors. Bottom plot: even order with errors.

This process exposed several issues in code relating to handling of longitudinal dynamics and spin precessing vertical and horizontal kickers.

Table 1: Summary of Spin Tune Check

Case No.	ϕ Snake	θ Snake	Hor. orbit angle at Snake	$\Delta p/p$	Q_s Zgoubi	Q_s TEAPOT-SPINK	Q_s Analytical
1	no snake	–	0.0	0.0	0.267	0.267	0.267
2	180	90	0.0	0.0	0.4997	0.5	0.5
3	180	90	0.2375	0.0	0.471	–	0.471
4	180	90	0.218	0.0	–	0.474	0.474
5	180	90	0.0	0.001	0.494	0.4955	0.4952
6	180	85	0.0	0.0	–	0.472	0.4722

IMPACT OF SLICING

It is well known that establishing the correct amount of slicing is critical for accurate spin transport. Previous work had indicated 15 slices per quad or approximately 0.2 m slice per ratio necessary to achieve convergence using old thick orbit element tracking with SPINK below γ of 100. As should have been intuitively obvious we found that this ratio depended on the rate of dS/dt which is a function of energy (since it determines basic spin precessing frequency through dipoles) and the proximity to spin resonances.

In Fig. 3 we can see how at energies below 230 GeV and away from a snake resonance we find 8 slices which is equivalent to 16 slices in the old SPINK code is sufficient. However in Fig. 4 as we approach the 7/10 snake resonance this resolution needed to be increased beyond 30 slices per quad.

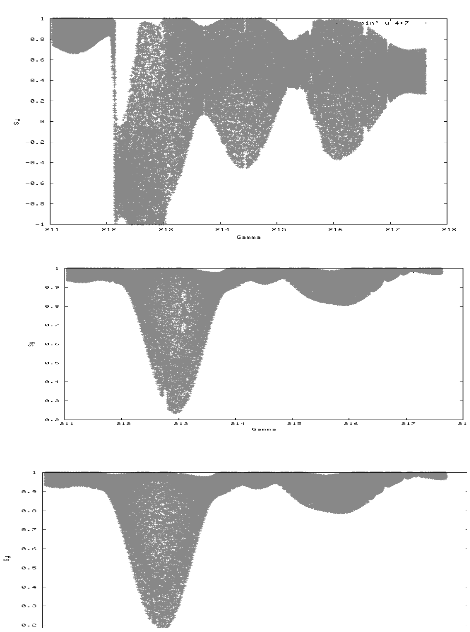

Figure 3: Four slices per quad (top), 8 slices per quad (middle), 200 slices per quad (bottom).

We later on discovered that both energy and acceleration rate can play a similar role in determining the appropriate

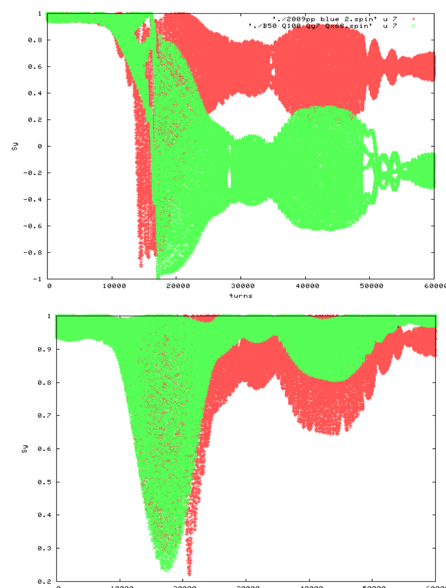

Figure 4: Top plot: (on resonance Qy=0.70) red trace 30 slices per quad, green 200 slices. Bottom plot: (close to resonance) right plot Qy=0.688 red trace 8 slices per quad, green 200 slices.

slices. During tracking above 230 GeV using 8 slices under realistic acceleration rates we found non-physical depolarizations around the last two strong intrinsic spin resonances in the 250 GeV acceleration ramp with small orbit errors. Comparisons with Zgoubi and previous spin tracking studies revealed this to be non-physical. Indeed when we increased the slicing from 8 to 30 slices this depolarization disappeared under 10 times faster acceleration rate. However tracking again with realistic acceleration rates showed that even this was not enough and we found it necessary to boost the slice up to 64 slices for the large triplet quadrupoles. The increased slices would have increased the tracking time necessary by 8 times. However removing orbit and spin transparent elements combined with targeted hand slicing of just the triplet quads kept the increase to only 2 times slower.

PORTING TEAPOT-SPINK TO GPU

Currently we are interested in studying from 100–250 GeV acceleration region in RHIC. If we slice at a ratio of every 0.1 m of quadrupole field up to 230 GeV and then restart using 0.05 m from 230 GeV to 250 GeV it will cost a total of 58 hours on the NERSC's Carver machine. We can run with a maximum 128 cores for 48 hours, which means we can track only 128 particles per que submission. The use of GPU could potentially permit 1000s of particles per core to be tracked albeit at a slower rate. So we have recently ported the TEAPOT-SPINK algorithm described in this paper to a GPU.

We initially attempted to develop separate GPU integrators for orbit using TEAPOT and for spin using SPINK as was done in the UAL version of SPINK which called the TEAPOT library for orbit pushes in between sliced Spin kicks.

However we found this approach to be very difficult, clumsy and slow to accomplish in CUDA code since CUDA does not yet lend itself to a true object oriented C++ model. So we took the TEAPOT orbit push algorithm and re-wrote it inside of the UAL library version of SPINK. Now the Cuda part of the code was all contained in only in the SPINK UAL library. This version of the code was able to achieve 20 fold speed pushing 100K particles over serial CPU execution (2.4 GHz Intel 5530 node) . However even at only 200 particles per turn (the threshold number of particles when GPUs will out perform the serial CPU execution) the time per turn was 2.5 s which is too slow to be able to scan through Spin resonances of interest which we would need 500,000 turns at realistic acceleration rate or ∼14 days of tracking on the NERSC GPU machine under double precision (NVIDIA Tesla C2050 code named Fermi).

Part of the problem was that we were keeping the lattice data on the CPU host memory and transferring it to the GPU as needed and as well the structure of our code involved many individual calls to the GPU kernels to propagate the orbit and spin vectors. This incurred large latency transferring memory back and forth between the CPU and GPU. So we re-wrote the cuda code to load the whole lattice into GPU global memory and placed the particle data in the GPU registers (the fastest memory on a GPU). All calls to perform the various kicks, were placed from inside the GPU to 'device' kernels. Now the CPU only needed to place a single call to the GPU returning periodically dump spin and orbit data. In Fig. 5 the comparison in time per turn between CPU and GPU Fermi machine on the DIRAC cluster at Nersc is shown. We see that while below 100 particles CPU generally out performs the GPU above 100 particles ratio can be as high as 20 times as one reaches 100K particles. But for the purposes of spin tracking the time per turn for 100K particles is far too slow to produce meaningful physics results. We are currently focused on the range of 1000 particle per GPU which at 0.7 s/turn costs 97 hours for 500K turns. If we were to use all 32 GPU's per node available under a regular que we could push 32,000 particles at a realistic acceleration rate across the last strong intrinsic resonance location in 97 hours. This is still a long time but clearly out performs what would be possible using conventional MPI - CPU approach.

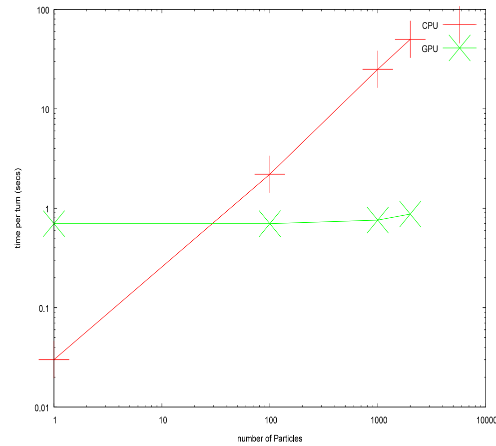

Figure 5: Log plot of time per turn through RHIC lattice for CPU (red) and GPU (green).

CONCLUSION

We have benchmarked and corrected several issues of new TEAPOT-SPINK code in the UAL framework. In addition we have ported a self contained version of TEAPOT-SPINK to the GPU platform. The performance now is at a stage where we can begin to consider isolated resonance crossings using realistic particle distributions. However more work needs to be performed to try optimize the performance of the cuda code. Several areas yet to be explored are: 1) Moving the lattice data to the registers memory location on the GPU 2) Consolidating drift and other lattice elements 3) more judicious use of double precision versus float (floating point calculations are x2 faster than double on the Fermi machines).

ACKNOWLEDGEMENTS

This work was funded by DOE/NP Grant No. DE-SC0004432 and Tech-X Corp. We would also like to thank NERSC for use of their computing resources.

REFERENCES

[1] L. Schachinger, R. Talman, "Teapot: A Thin Element Accelerator Program for Optics And Tracking." Part. Accel. 22:35, 1987

[2] E.D. Courant and R.D. Ruth, BNL Report No. 51270, 1980

[3] A. Luccio, BNL-52481, BNL-AGS-AD-95-2, Sep 1995.

[4] http://mad.home.cern.ch/mad/mad8web/mad8.html

[5] Francois Meot and S. Valero, FERMILAB-TM-2010, DAPNIA-SEA-97-13, Oct 1997.

[6] N. Malitsky et al., Proceedings of Particle Accelerator Conference (PAC 03), Portland, OR, 12-16 May 2003, pp 3482.