SLAC-PUB-10577
July, 2004

Start to End Simulations of Low Emittance Tuning and Stabilization[1]

P. TENENBAUM, A. SERYI, M. WOODLEY
Stanford Linear Accelerator Center, Stanford University, Stanford, CA
D. SCHULTE
CERN, Geneva, CH
N. WALKER
DESY, Hamburg, DE
G. WHITE
Queen Mary University, London, UK

Abstract

The principal beam dynamics challenge to the subsystems between the damping ring and the collision point of future linear colliders is expected to be the tuning and stabilization required to preserve the transverse emittance and to collide nanometer-scale beams. Recent efforts have focused on realistically modelling the operation and tuning of this region, dubbed the Low Emittance Transport (LET). We report on the development of simulation codes which permit integrated simulation of this complex region, and on early results of these simulations. Future directions of LET simulation are also revealed.

Invited talk at *The Ninth European Particle Accelerator Conference,* Lucerne, Switzerland, 05–09 July 2004.

[1]Work supported by the Department of Energy, Contract DE-AC03-76SF00515.

START TO END SIMULATIONS OF LOW EMITTANCE TUNING AND STABILIZATION†

P. Tenenbaum‡, A. Seryi, M. Woodley, SLAC, Menlo Park, CA USA
D. Schulte, CERN, Geneva CH
N. Walker, DESY, Hamburg, DE
G. White, Queen Mary University, London, UK

Abstract

The principal beam dynamics challenge to the subsystems between the damping ring and the collision point of future linear colliders is expected to be the tuning and stabilization required to preserve the transverse emittance and to collide nanometer-scale beams. Recent efforts have focused on realistically modelling the operation and tuning of this region, dubbed the Low Emittance Transport (LET). We report on the development of simulation codes which permit integrated simulation of this complex region, and on early results of these simulations. Future directions of LET simulation are also revealed.

INTRODUCTION

In order to achieve an acceptable particle physics event rate in the face of cross-sections which scale as $1/E^2_{\rm CM}$, it will be necessary for a future linear collider operating at 500 GeV CM to achieve a luminosity significantly in excess of $10^{34}\ {\rm cm}^{-2}{\rm s}^{-1}$. At this time, there are three main candidate designs for future linear colliders: TESLA, which uses superconducting RF cavities at a relatively low frequency of 1.3 GHz [1]; GLC/NLC, which uses room-temperature RF structures at a relatively high frequency of 11.4 GHz [2]; and CLIC, which uses room-temperature RF structures at a very high frequency of 30 GHz, which are excited by a high-power, low energy drive beam instead of conventional RF sources such as klystrons [3]. Despite the design differences which are driven by the choice of main linac technology, all three candidate linear colliders use identical strategies to achieve their luminosity goals: long trains of bunches are stored in damping rings which reduce their vertical normalized emittances to approximately 10 nm.rad; after extraction from the damping ring the bunches in the train are compressed longitudinally by a factor of 20-50; the trains of short bunches are accelerated from a few GeV to 250 GeV in a long linear accelerator; the beams are demagnified to RMS vertical sizes of a few nanometers in a chromatically-corrected beam delivery system and collided. Table 1 compares a few relevant parameters for the three designs at 500 GeV CM. Note that while the acceleration technologies and beam time structures (bunches per train and trains per second) are wildly different between the three designs, the single-bunch parameters are generally similar.

† Work supported by the U.S. Department of Energy, contractDE-AC03-76SF00515.

‡ quarkpt@slac.stanford.edu

Table 1: Key parameters for 3 linear collider designs (taken from [4] except as noted).

		TESLA	GLC/NLC	CLIC
Bunch Population	10^{10}	2	0.75	0.4
Bunches/ Train		2820	192	154
Bunch Separation	ns	337	1.4	0.67
Repetition Rate	Hz	5	120	200
Damping Ring				
Energy	GeV	5.0	1.98	2.42
$\gamma\epsilon_y$	nm	20	20	5
σ_z	mm	6	5.5[5]	1.3
Collision Point (250 GeV/beam)				
$\gamma\epsilon_y$	nm	30	40	10
β^*_y	μm	400	110	50
σ^*_y	nm	5	3	1.5
σ_z	μm	300	110	35
$\mathcal{L}$	$10^{34}{\rm cm}^{-2}{\rm s}^{-1}$	3	2	2

The region from the extraction point of the main damping ring to the collision point is collectively referred to as the Low Emittance Transport (LET). The LET was investigated by the International Linear Collider Technical Review Committee (ILC-TRC) as part of its second study on the state of the art, published in 2003 [4]. That study found that the fundamental design of each LET, in the absence of errors or misalignments, is sound and can deliver the desired luminosity. Therefore, the obstacles to luminosity production are universally associated with static or dynamic imperfections of the LET implementation which must be addressed through tuning and stabilization.

Since the achievable luminosity is such a strong function of the performance of the tuning and stabilization algorithms, it is necessary to be able to reasonably estimate their performance. Experience has shown that it is almost never possible to derive an analytic expression for the performance of a complicated tuning algorithm, and such expressions usually require drastic simplifications of the problem. Given this state of affairs, the luminosity performance of a linear collider LET can only be estimated by direct simulations which thoroughly emulate the tuning procedure and include all errors and performance limitations which will be encountered in real-world operations.

ISOLATED VS. INTEGRATED SIMULATIONS

During the design phase of the LET, the main subsystems (bunch compressor, main linac, beam delivery system) were generally developed in isolation from one another. All of the relevant parameters were specified at any interface between subsystems to ensure that the designs did not diverge from one another. As part of this process, simulation studies of tuning and stabilization were typically performed on each major subsystem in isolation, and these studies have been reported in a variety of venues, as summarized in the TRC report of 2003.

Such an approach is perfectly acceptable for studying a wide class of errors in which the effects can be combined by simple rules. For example, quadrupole misalignments produce growth in RMS emittance which is linearly additive; designers simulating the beam-based alignment of several subsystems can simply add the normalized emittance growths from this source in each subsystem to compute a total growth. Similarly, beam jitter driven by quadrupole vibration can be summed in quadrature to compute a total beam jitter for comparison to a jitter budget.

The circumstances under which isolated simulations become unsatisfactory are twofold. The first circumstance is errors which are not simple to combine. As an example, beam delivery systems often include complicated chromogeometric aberrations. The emittance growth generated by these aberrations is not a fixed quantity but depends on the incoming emittance, therefore changing the emittance growth generated in the main linac can also change the amount generated in the beam delivery. The second circumstance is when an error in one subsystem can change the performance of a tuning algorithm in a downstream subsystem. This is particularly important in the LET of a linear collider, where beam signals are used by virtually every tuning algorithm of consequence to the luminosity performance. As an example, tuning errors in the bunch compressor can lead to an incorrect longitudinal phase relationship between the damping ring extraction and the IP; in such a case the RMS bunch length and energy spread of the beam at the exit of the bunch compressor can be correct, but the resulting transformation of damping ring phase jitter to IP energy jitter can impede tuning of the beam delivery system.

A further impetus to the development of integrated simulations is the intensity of the beam-beam interaction. In the TESLA design, for example, the beam-beam interaction is so intense that longitudinally-correlated distortions of the transverse shape of a bunch which are too small to produce significant growth in the RMS beam emittance can nonetheless cause substantial reduction in the luminosity [6]. In addition, the beam-beam interaction causes the luminosity loss from an offset at the collision point to vary significantly from what would be predicted for Gaussian beams in the absence of the beam-beam interaction, as shown in Figure 1.

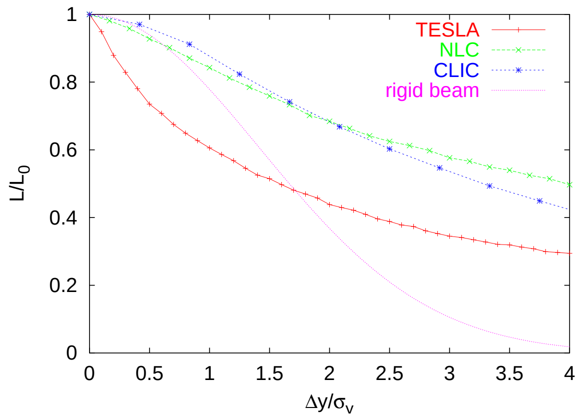

Figure 1: Luminosity as a function of beam-beam offset at the collision point for TESLA, GLC/NLC, and CLIC, compared to rigid Gaussian beams which do not experience beam-beam interaction.

In summary, then: while isolated simulations of the LET subsystems are adequate for the design phase, it is generally not sufficient for estimating the luminosity performance of the LET as a whole once the systems are connected. For this it is necessary to simulate the two LETs of the complex in an integrated fashion, including the beam-beam interaction at the collision point. It is also worthwhile to note that the end result of an integrated simulation of the tuning and operation of the LET is a more realistic model of the actual machine which can be used for additional studies. For example, such a post-tuning model of the LET is potentially a more accurate predictor of detector backgrounds and collimation efficiency than the error-free "perfect" LET model.

PERFORMANCE ISSUES

The main reason that fully integrated simulations have not previously been undertaken for linear colliders is execution speed. Although modern computers are vastly more powerful than the models used to design the Stanford Linear Collider and the Final Focus Test Beam at SLAC, the size and complexity of linear collider LETs still present obstacles to completing integrated simulations in a timely manner.

Lattice Representation

The region between the damping ring and the collision point in a linear collider can contain thousands of magnets, thousands of beam position monitors, and tens of thousands of accelerating structures. A fully-instantiated data structure for the beamline, including information on fixed relationships between the elements (for example, keeping track of multiple magnets powered in series or supported on a common girder) can require as much as 100 megabytes of memory. Since complete simulations require two full LETs, the memory consumption becomes significant.

Accelerator simulation codes written in earlier times resisted full instantiation in order to limit memory consumption. Modern LET codes have moved very strongly in the direction of full instantiation in order to more accurately model reality, since every element is unique and has a unique set of misalignments and errors. Although memory consumption of 200 MB or more is not unacceptable by modern standards, the performance of a simulation which must manipulate such a large data space must be taken into account when designing the data structures which represent the accelerator.

Beam Representation

In a linear collider LET, the most common element class is typically the RF structure which provides the linear acceleration. The representation of the beam which is used in tracking should therefore be selected to optimize the tracking speed through an RF structure.

The most CPU-intensive operation in tracking through an RF structure is applying the transverse wakefield from leading particles to trailing particles. The computation time required scales with the square of the number of particles, and therefore the key to speed is reducing the number of particles to a tolerable level. For linac codes this is typically accomplished by dividing each bunch longitudinally into slices, and at each slice position representing the beam with a small number of macroparticles. In addition to a fixed z position each macroparticle has an energy, a 4-dimensional transverse centroid, and a 4 x 4 matrix of second moments in the transverse. The incoherent energy spread is represented by assigning a different energy to the macroparticles at a given z location. This representation is ideal for the application of transverse wakefields, and is also appropriate for the application of longitudinal wakefields and linear acceleration (which varies sinusoidally with longitudinal position in the bunch). Since the macroparticles include transverse degrees of freedom up to the second moment, the transverse optics of structures, drifts, and quadrupoles can be included as well. By using this representation a simulation can represent a bunch with a few hundred macroparticles with good accuracy.

The macroparticle approach has a number of drawbacks when simulation of the rest of the LET is considered. The momentum compaction of bending magnets in the bunch compressor will cause macroparticles to migrate in z; when this happens the representation of incoherent energy spread becomes more complicated, and the total number of slices will typically increase, leading to less efficient application of wakefields. Beams in which the transverse degrees of freedom are included only up to second moment are not adequate for simulation of beam delivery systems, which always include sextupoles and often include octupoles and other high-order multipoles. Such a representation is often inadequate in the fringe field of bend magnets as well. Finally, modelling the luminosity generated in the beam-beam interaction also requires that higher-order transverse moments be known, since it is well known that RMS beam size is a poor predictor of luminosity for beams which deviate significantly from Gaussian transverse distributions. Under these circumstances the most accurate way to represent the beam is with a large ensemble of dimensionless rays, which are initially randomly positioned in such a way as to emulate the beam's initial distribution (usually Gaussian in 6 degrees of freedom).

We see, therefore, that the beam in a simulation of the LET must at certain times have the properties of a macroparticle representation while at other times must have the properties of a dimensionless ray representation.

Multiple Bunches

Each linear collider design requires trains of bunches to achieve its design luminosity. Simulation of multiple bunches during one tracking operation will typically extend the time needed for tracking in proportion to the number of bunches per train, although there are some circumstances under which an even less favorable slowdown can be arranged.

Most tuning and stabilization studies can be carried out without use of bunch trains. Indeed, this is appropriate since most tuning operations will be performed in single-bunch mode. Stabilization studies can under some circumstances be performed in single-bunch mode, but sometimes it is necessary to use at least a partial bunch train. For example, the TESLA design includes several feedback loops that operate within a single bunch train [7], which is best studied with an actual train and with intra-train tuning operations enabled in simulation.

Multibunch tracking code requires that some care be taken with the design choices based on the expected use of the code. For example, a code which uses an outer loop to loop over bunches and an inner loop over elements will be less efficient than a code which loops in the opposite order, since the latter code needs to preserve long-range wakefield kick information for only one element at a time. Use of an outer bunch loop facilitates intra-train tuning, which is not straightforward when an outer element loop is used.

EXISTING CODES

In recent years several simulation tools have been used to study the full LET of a linear collider. None of the tools in question was originally intended for such a purpose; each one was adapted to allow it to simulate the LET. We review the most salient examples below.

MAD-8 [8]: MAD-8 includes high-order and path-length calculations of a wide variety of magnet classes as well as tracking of rays, but does not include linear acceleration. A version of MAD-8 was adapted to include linear acceleration and linac wakefields [9]. This program has been used for linac studies at the design stage but is not generally used for tuning studies.

PLACET [10]: PLACET was written to study beam dynamics and tuning in the CLIC drive beam and main linac. PLACET permits either rays or macroparticles to be tracked, and was extended to study final focus systems with dispersive bend magnets and thin-lens sextupoles or higher multipoles. PLACET does not treat the momentum compaction of bend magnets, but was adapted to study bunch compressors by running MAD as a subprocess, since MAD does properly treat momentum compaction.

LIAR [11]: originally written solely for studying the performance and tuning of linacs, LIAR uses a macroparticle beam representation. LIAR was extended to study bunch compressors and beam delivery systems through integrating the DIMAD [12] tracking engine into LIAR. Since DIMAD represents the beam as rays, it is necessary to transform between rays and macroparticles several times in the course of tracking the LET.

MERLIN [13]: A library of accelerator physics routines originally written to study performance and tuning of beam delivery systems. MERLIN was completely rewritten for more general use in all LET regions, and now supports both ray and macroparticle tracking.

GUINEA-PIG [14]: a program which simulates the beam-beam interaction and estimates luminosity and deflection as well as effects of interest to particle physicists (pair production, etc.). GUINEA-PIG is the standard tool for linear collider beam-beam interactions; all simulations which include beam-beam effects do so by interfacing an LET tracking code with GUINEA-PIG.

SAMPLE STUDIES

At this time, the holy grail of LET studies – a simulation in which all key beam parameters have been tuned up via beam-based signals and algorithms applied from the bunch compressor to the collision point – remains unrealized. Several studies have been performed which moved in that direction, and we report on two of them here.

Luminosity in the Presence of Ground Motion

Given that all linear collider designs seek to collide beams with RMS vertical sizes in the nanometer regime, it is essential to ensure that the beams can be kept in collision in the presence of natural and artificial vibration sources. One of the key tools for maintaining collisions is a steering feedback loop which uses the beam-beam deflection to measure the offset of the beams at the collision point, and this deflection is known to be sensitive to the longitudinally-correlated beam distortions which are introduced by wakefields in the main linac. In order to ensure that these distortions accurately mimic those expected in real life, it was necessary to track the beam through a beamline which had been misaligned and tuned in such a way as to produce both the correct DR-to-IR emittance dilution and the correct bunch shape distortions. Other prerequisites include: one or more plausible models for natural ground motion, plausible models for detector motion (since particle physics detectors often generate substantial vibrations), and reasonable feedback and correction algorithms.

The study of ground motion in a realistic setting was reported previously [15], and we summarize the procedure and results here. The study was performed by first applying *ab initio* misalignments to the main linac and then using steering algorithms to tune them to the IP emittances specified in Table 1; although the beam was tracked through the main linac, the beam delivery, and the bunch compressors in the case of GLC/NLC and TESLA, only the main linac was misaligned and steered. Two LETs were prepared in this way, and the resulting beams brought into collision. At that point the time-averaged luminosity was studied as a function of ground motion model (from quiet site model “A” to noisy site model “C”), presence or absence of additional detector motion, final doublet mechanical stabilization, and several different types of IP beam-beam feedback. Figure 2 summarizes the results; interested readers are strongly encouraged to read a more detailed discussion of the study in [4, 15].

Use of Intra-Train Optimization

The study discussed above used idealizations which did not yield a full understanding of feedbacks which operate within a single TESLA bunch train. Of particular interest was the speed and accuracy with which an intra-train optimization of the collision offset and beam-beam angle would converge on the point of maximum luminosity. Such a simulation demanded all the prerequisites of the vibration study (save those specific to vibration), plus the long-range transverse wakefields of the TESLA accelerating structures.

Figure 3 shows the luminosity as a function of bunch number when a realistic optimization of the luminosity is included. The initial fast rise from zero luminosity is due to the action of the beam-beam feedback, which quickly minimizes the centroid deflections; the structures around bunch 150 and bunch 300 are due to collision offset and collision angle optimization scans, in which the parameter of interest is scanned and the luminosity measured in order to find an optimum. In this case, GUINEA-PIG is used to track e^+e^- pairs produced in the collision to the detector luminosity monitor in order to include the luminosity measurement process in the simulation.

FUTURE DIRECTIONS

One of the most glaring features of the list of existing simulation tools is that none of them was explicitly designed for the purpose of simulating the full LET of a linear collider. Instead, each program was written for a different purpose and later adapted to the task of LET simulation, with varying degrees of success. Conceptual work has begun on the design of a simulation tool which is more thor-

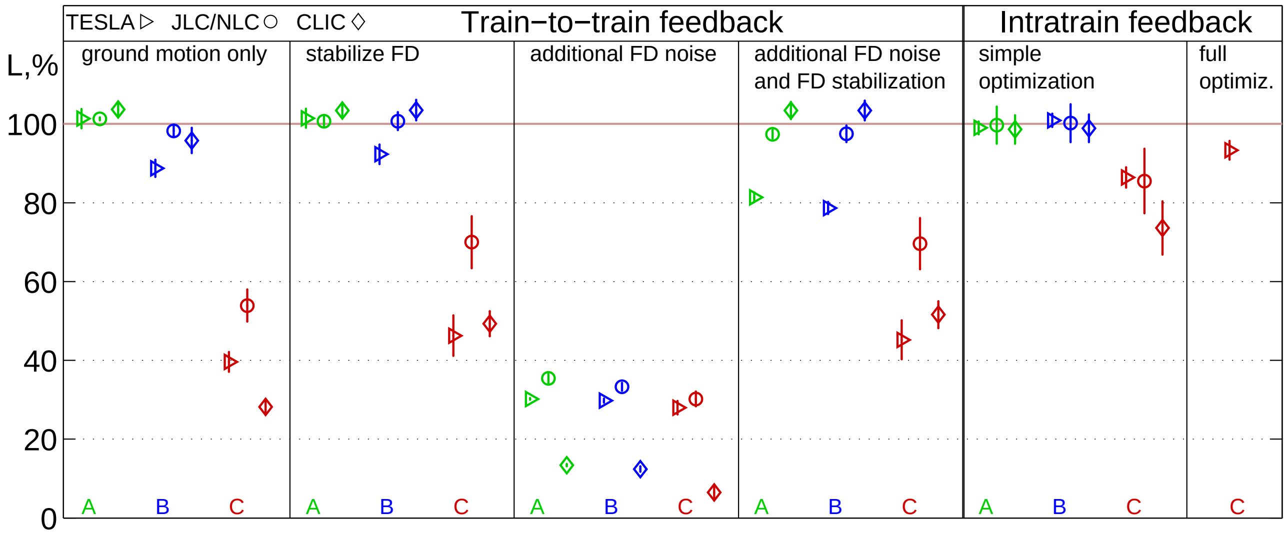

Figure 2: Results of integrated study of vibration and stabilization on mean luminosity.

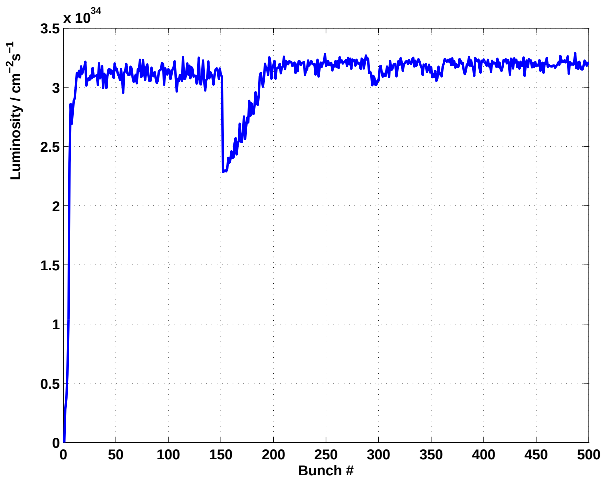

Figure 3: Luminosity as a function of bunch number during a TESLA IP optimization scan.

oughly optimized for the task of complete simulation [16]. Such a tool would be useful not only for linear colliders but also for linac-based light sources, which combine linear acceleration, sophisticated bunch compressors, and beam delivery systems with tight tolerances.

On the algorithm side, the longitudinal tuning requirements of the bunch compressors have been neglected, in part due to a greater perceived urgency to the study of transverse dynamics in general and emittance preservation in the main linac in particular. Strong and rapid progress on the latter in recent years implies that it will soon be possible to more thoroughly integrate the existing and future tuning studies into a single simulation.

REFERENCES

[1] *TESLA Technical Design Report* (2001).

[2] P. Tenenbaum, "The JLC/NLC Baseline Design," Proc. PAC-03 (2003).

[3] G. Guignard, ed., "A 3 TeV e+e- Linear Collider Based on CLIC Technology," CERN-2000-008 (2000).

[4] *International Linear Collider Technical Review Committee Second Report* (2003).

[5] M. Woodley and A. Wolski, "The NLC Main Damping Ring Lattice" (2003).

[6] R. Brinkmann, O. Napoly, D. Schulte, "Beam-beam Instability Driven by Wakefield Effects in Linear Colliders," proc. PAC-01 (2001).

[7] R.-D. Kohaupt, I. Reyzl, "Fast Feedback Systems for Orbit Correction in the TESLA Linear Collider," Proc. PAC-99 (1999).

[8] H. Grote *et al*, "The MAD Program," Proc. PAC-89 (1989).

[9] H. Grote *et al*, "Extension of MAD Version 8 to Include Beam Acceleration," Proc. EPAC-00 (2000).

[10] D. Schulte, "PLACET: A Program to Simulate Drive Beams," Proc. EPAC-00 (2000).

[11] R. Assmann *et al*, "LIAR: A Computer Program for the Modelling and Simulation of High Performance Linacs" (1997).

[12] R. Servranckx *et al*, SLAC-285 (1995).

[13] `http://www.desy.de/ merlin`.

[14] D. Schulte, CERN-PS-99-14 (1999).

[15] A. Seryi *et al*, "Effects of Dynamic Misalignments and Feedback Performance on Luminosity Stability in Linear Colliders," Proceedings PAC-03 (2003).

[16] `http://www.slac.stanford.edu/~quarkpt/lucretia`.