SLAC–PUB–5355
December 1990
(A)

A NEW TECHNIQUE OF CORRECTING EMMITTANCE DILUTIONS IN LINEAR COLLIDERS *

T. O. RAUBENHEIMER

Stanford Linear Accelerator Center, Stanford University, Stanford, CA 94309 USA

Abstract

This paper describes a new method of reducing the transverse emittance dilution in linear colliders due to both transverse wakefields and dispersive errors. The technique is a generalization of the Dispersion-Free [1] correction algorithm; the dilutions are corrected locally by varying the beam trajectory. This technique will complement the BNS damping [2] method which primarily corrects the dilutions resulting from coherent betatron oscillations. Finally, the results of simulations are presented demonstrating the viability of the technique.

Submitted to Nuclear Instruments and Methods A

---

*Work supported by Department of Energy contract DE–AC03–76SF00515.

1. Introduction

In a linear collider the magnets, accelerating structures, and the Beam Position Monitors (BPMs) are all typically misaligned relative to the ideal centerline. Thus, the beam trajectory is offset in both the magnets and the accelerating structures. This can lead to transverse wakefields and dispersive errors which dilute the beam's (projected) transverse emittance, a measure of the phase-space volume occupied by the beam. These dilutions then cause a reduction in the collider's luminosity. This paper will describe a new technique of correcting these two sources of transverse emittance dilution.

Both of these dilutions depend upon the transverse alignment of the accelerator relative to the beam size. To achieve the necessary luminosity in future linear colliders the beam sizes are very small, and, if uncorrected the wakefields and the dispersive errors would impose extremely tight transverse alignment tolerances on the collider. Thus, correction of these dilutions is crucial for future linear collider designs.

A particle beam consists of particles distributed in six-dimensional phase-space: $(x, p_x, y, p_y, z, \Delta E)$ where the first four coordinates specify the position in the two transverse planes and z and ΔE are the longitudinal position in the bunch and the energy deviation. Obviously in a conservative system, which a high-energy linear accelerator approximates, the total phase volume is conserved. However, the luminosity is strongly dependant upon the *projection* of this phase volume onto the transverse planes. Since the transverse and the longitudinal degrees-of-freedom are initially uncoupled, any added correlations will increase this projected emittance thereby decreasing the luminosity.

Both the transverse wakefields and the dispersive errors do just this; they correlate the longitudinal and the transverse degrees-of-freedom. Transverse wakefields result from the electromagnetic interaction between the particle bunch and its surroundings, namely, the acceleration structures. When a point charge travels off-axis in a structure, it leaves behind a transverse wakefield that will deflect subsequent particles. These deflections cause a particle's trajectory to be a function of it's longitudinal position within the bunch and thereby cause a dilution of the (projected) transverse emittance. Likewise, dispersive errors arise when the beam travels off-axis. If the beam is offset in a quadrupole magnet, it will be deflected. Since particles with different energies are deflected differently, the trajectory will be a function of the energy deviation and the transverse emittance will be diluted.

2. Theory

The equation of motion for a particle in a high-energy linear accelerator can be written [3]

$$\begin{aligned}\frac{1}{\gamma(s)}\frac{d}{ds}\,\gamma(s)\,\frac{d}{ds}\,x(s;z,\delta) + (1-\delta)\,K[x(s;z,\delta)-x_q] = (1-\delta)\,G \\ + \frac{(1-\delta)}{\gamma_0(s)}\,Nr_0\int_z^{\infty} dz' \int_{-\infty}^{\infty} d\delta' \rho(z',\delta')\,W_\perp(s;z-z')\,[x(s;z',\delta')-x_a] \quad ,\end{aligned} \tag{2.1}$$

where s and z are the longitudinal position in the accelerator and in the bunch, and δ is the relative energy deviation: $\delta = \big(\gamma(s)-\gamma_0(s)\big)/\gamma(s)$; note that the energy deviation is also a function of s and z. Next, K and G are the normalized focusing and bending functions: $K(s) = \frac{e}{p_0 c}\frac{dB_y}{dx}$ and $G(s) = \frac{e}{p_0 c}B_y$, where p_0 is the design particle momentum and B_y is the vertical component of the magnetic field. N and r_0 are the number of particles and the classical electron radius, $W_\perp$ and ρ are the transverse wakefield and the longitudinal distribution function for the particle bunch, and finally, x_q and x_a are the misalignments of the quadrupoles and the accelerator structures.

As one can see in eq. (2.1), the transverse position is coupled to both longitudinal coordinates; our goal is to minimize the effect of this coupling. We can gain some insight into the problem by using two two-particle models: one to determine the wakefield effects and one for the dispersive effects. First, consider two macro-particles, each with charge $N/2$, located at $z = \pm\sigma_z$, where σ_z is the rms bunch length. To determine the effect of the wakefields, we examine the difference between the trajectories of these two macro-particles $\Delta x_w = x(\sigma_z, 0) - x(-\sigma_z, \overline{\delta})$ where the energy deviation of the second particle $\overline{\delta}$ is the "correlated" energy spread; this is an energy spread which is correlated with z. Next, to find the effect of the dispersive errors, we consider the difference between the trajectory of the on-energy head particle and a particle, also located at $z = \sigma_z$, with an "uncorrelated" energy spread ξ: $\Delta x_d = x(\sigma_z, 0) - x(\sigma_z, \xi)$.

Assuming that the wakefields and ξ are small, we can use eq. (2.1) to solve for Δx_w and Δx_d perturbatively. The first order solutions are

$$\begin{aligned}\Delta x_w(s) = \int_0^s ds' R_{12}(s,s') \Big[\overline{\delta}\Big(G + Kx_q\Big) \\ - \Big(\overline{\delta}K + \frac{Nr_0}{2\gamma_0} W_\perp(2\sigma_z)\Big)x + \frac{Nr_0}{2\gamma_0} W_\perp(2\sigma_z) x_a \Big] \quad ,\end{aligned} \tag{2.2}$$

and

$$\Delta x_d(s) = \int_0^s ds' R_{12}(s,s') \Big[\xi\Big(G + Kx_q\Big) - \xi K x \Big] \quad , \tag{2.3}$$

where x is the trajectory of the on-energy head particle $x(\sigma_z, 0)$ and $R_{12}(s, s')$ relates a deflection at s' to a position as s; the R_{12} is the Green's function for the focusing structure of the accelerator. The first term (enclosed in parentheses) in both equations will tend to be small. This occurs when the beam trajectory is corrected since the dipole correctors G are adjusted to cancel the quadrupoles deflections Kx_q. Since the correction is performed locally, keeping the trajectory offsets small throughout the machine, the cancellation is independent of slow variations in $\overline{\delta}$ and ξ. Thus, the primary sources of emittance dilution are the last two terms of eq. (2.2) and the last term of eq. (2.3).

Looking at eqs. (2.2) and (2.3), one sees that there are two free parameters that can be varied to correct the dilutions: the correlated energy spread $\overline{\delta}$ and the trajectory $x(s)$. The BNS damping technique [2] does the former; it uses $\overline{\delta}$ to reduce the second term of eq. (2.2). Specifically, $\overline{\delta}$ is adjusted so that $\overline{\delta}K$ cancels the wakefield $Nr_0W_\perp/2\gamma$. In the smooth approximation, where K and $W_\perp$ are smooth functions of s, one can solve for a $\overline{\delta}$ such that this term is always zero. Unfortunately, this local cancellation is not possible* in the alternating-gradient focusing structures used in high-energy machines. While the wakefield $W_\perp$ has a constant sign, an alternating-gradient focusing structure usually contains a periodic array of discrete focusing magnets with both positive and negative K values. Since the energy spread $\overline{\delta}$ cannot be changed rapidly with s, at best one can adjust $\overline{\delta}$ to cancel the integral of this term over a cell of the focusing structure. Furthermore, since this cancellation depends upon the position x in the quadrupoles and the accelerator sections, exact cancellation is only possible if $x(s)$ is correlated from point-to-point. This is the case for a coherent betatron oscillation, but it is not true if the particle is steered or deflected by random errors as is the case for a corrected trajectory. Thus, while the BNS technique can cancel the wakefield effects due to

---

* It was assumed here that $\overline{\delta}$ is due to an energy deviation. It is also possible to vary the focusing strength with RF quadrupoles.

a coherent betatron oscillation, it may reduce, but cannot cancel, the effects of wakefields due to a corrected trajectory.

The Dispersion-Free (DF) [1] correction technique uses the later approach to correct the dispersive emittance dilution. Here, the trajectory $x(s)$ is varied so that over any short region of the accelerator the integral in eq. (2.3) is small. The technique "measures" the dispersive errors by measuring the difference of two trajectories while changing the beam energy, or equivalently, while changing the magnet strengths. The equation for this difference orbit is identical to eq. (2.3) except that ξ is replaced by the effective energy change; this is typically around 10%. Thus, the difference orbit will accurately reflect the emittance dilution except for measurement errors and effects of the non-linearity of the dispersive error; a complete analysis of all the errors in given in ref. 1. By correcting the difference orbit, in concert with the actual trajectory, the DF correction technique can reduce the emittance dilution to negligible values.

3. Wake-free correction

Given the performance of the DF algorithm, we have attempted to extend it to also correct wakefields. The goal is to find a new trajectory along which both the wakefield and the dispersive effects cancel. The wakefields are caused by trajectory offsets in the accelerator sections which are due to both misalignments of the accelerator sections and a non-zero trajectory. If we ignore the accelerator misalignments, the effective offset in a section is just the average of the position in the two adjacent quadrupoles. By varying the quadrupole strengths in a specified manner, one can measure a difference orbit where the orbit in the quadrupoles will *mimic* the effects of the wakefields due to the trajectory. From the second term in eq. (2.2), we find that, to mimic the wakefield effect, the quadrupole strengths must vary as

$$\frac{\Delta K}{K} \propto \frac{L_{\mathrm{acc}}(s)}{\gamma(s)\, K L_{\mathrm{quad}}(s)} \sqrt{\frac{\beta_{\mathrm{acc}}}{\beta_{\mathrm{quad}}}} \cos\Delta\phi \quad , \tag{3.1}$$

where β_{acc} and β_{quad} are the beta functions at the middle of the accelerator sections and the adjacent quadrupoles and $\Delta\phi$ is the betatron phase advance between the two [3]. In addition, L_{acc} and L_{quad} are the lengths of the accelerator sections and the quadrupoles. Finally, note that because the correction is local, this condition can fluctuate slowly with s.

Condition (3.1) specifies that the quadrupole strength variation $\delta = \Delta K/K$ has opposite signs at focusing (QFs) and defocusing quadrupoles (QDs). In contrast, when creating the difference orbit to measure the dispersive error, δ has the same sign at both the QFs and the QDs. To correct both the wakefields and the dispersive errors one minimizes both of these difference orbits along with the actual trajectory. Unfortunately, it is not necessarily possible to increase some magnets while decreasing others since the quadrupoles are usually run close to their maximum strength. Thus, one can use an equivalent procedure where one minimizes a difference orbit created by varying only the QFs and a difference orbit created by varying only the QD magnets. In addition to being feasible, this later procedure also benefits from being simpler.

Strictly, by examining eq. (2.2), we can see that minimizing these two difference orbits will reduce the wakefields if the accelerator sections are aligned to the centers of the *quadrupoles*, not the machine centerline. This can be remedied by varying the dipole correctors when varying the quadrupoles. Thus, the dipole correctors (partially) cancel the effect of the quadrupole misalignments; they must or the trajectory would tend to grow. In practice, the correction technique works better if the accelerator structures are aligned to the quadrupoles, but, as will be demonstrated, it still works very well when the accelerator sections are aligned to the ideal centerline.

To recapitulate, the correction algorithm is: (1) measure a difference orbit $\Delta x_{QF}(s)$ created by varying the QFs and the associated dipole correctors, (2) measure the difference orbit $\Delta x_{QD}(s)$ created by varying the QDs and the associated dipole correctors, (3) measure the actual trajectory $x(s)$, and finally, (4) one minimizes all three of these orbits. When developing the DF algorithm, it was found that a weighted least-squares is the best minimization procedure. Thus, in this variation, one minimizes the sum:

$$\sum_{j\in\{BPM\}} \frac{(\Delta x_{QF})_j^2}{2\sigma_{\text{prec}}^2} + \frac{(\Delta x_{QD})_j^2}{2\sigma_{\text{prec}}^2} + \frac{x_j^2}{\sigma_{\text{BPM}}^2 + \sigma_{\text{prec}}^2} \quad ,$$

where each term is weighted by the accuracy of the respective measurement: σ_{BPM} is the estimated rms of the BPM misalignments and σ_{prec} is the rms precision (reading-to-reading jitter) of the BPM measurements. Although it does not correct the wakefields due to the accelerator section misalignments, this technique will be referred to as Wake-Free (WF) correction because the corrected trajectory does not cause wakefield or dispersive dilutions.

4. Simulations

In table 1, the performance of the DF and WF techniques is compared against a standard correction algorithm, the 1-to-1 method. The 1-to-1 algorithm adjusts the trajectory to zero the BPM measurements; typically, in this technique one only uses the BPMs and correctors located near the focusing quadrupoles. The correction was simulated in a preliminary design of the Next Linear Collider (NLC) [4] where the vertical beam size is tiny, roughly $2\,\mu$m. The results in table 1 are the average of correcting 20 sets of random error distributions and ϵ_{y0} is the initial undiluted vertical emittance. The error distributions have $70\,\mu$m rms vertical quadrupole and BPM misalignments, and $2\,\mu$m rms BPM precision errors; the accelerator sections were aligned to the ideal machine centerline. In addition, the optimal BNS energy spread has been added to the beam in all three cases. Finally, the initial conditions (y_0, y_0') were optimized [5] after 1-to-1 correction to further reduce the dilution; while this procedure reduces the dilution from nearly $50\,\epsilon_{y0}$ when using the 1-to-1 algorithm, it yields little improvement when using DF or WF correction. The WF technique performs extremely well; it virtually eliminates all of the emittance dilution and it does a better job correcting the trajectory than the other two methods.

Figure 1 shows plots of the beam distribution after (a) 1-to-1, (b) DF, and (c) WF correction for one of the 20 cases in table 1. The scatter-plots on the left are the projections of the beam distributions in the y-y' phase space while the right-hand plots are projections onto the y-z plane. One can immediately see that the beam emittance is seriously diluted after 1-to-1 correction; this is primarily due to dispersive errors. Next, after DF correction, the dispersive errors are corrected, but the distribution displays the tails characteristic of transverse wakefields; these arise from the random trajectory. Finally, after WF correction, one can see that the dilution due to both the dispersive errors and the wakefields is negligible.

5. Discussion

Before concluding, we need to discuss the correction of the wakefields due to accelerator section misalignments which the WF technique does not correct. One possible solution involves moving the accelerator sections or alternately using dispersion-free trajectory bumps [6]. If the higher-order terms of eq. (2.2) are negligible, then the correction is simple; one can correct the local errors with a global solution. Unfortunately, in the NLC and most other future linear collider designs, this is not true. In this case, one would like to correct the effects locally, but local correction requires local measurements; this is difficult. One method is a variation of the WF technique where one measures a difference orbit while varying the strength of the wakefield [7]. This can be accomplished by changing either the bunch population N or the bunch length σ_z. Another idea is to measure the wakefields directly by measuring the higher-order modes induced in the accelerating structures. This allows one to align the structures to the beam itself [8].

Barring local correction, one may still be able to eliminate the dilution by using multiple "knobs" or bumps to cancel the non-linear effects. When simulating 70 μm rms accelerator section misalignments in the NLC, we have used six dispersion-free bumps to reduce the emittance dilution from over 700% to less than 10%.

To conclude, we note that the WF technique reduces the emittance dilution due to misaligned quadrupoles and a non-zero trajectory extremely well. Since the technique is very similar to the DF method, we know that it is a robust algorithm and it is not sensitive to effects such as jitter and calibration errors. Furthermore, the technique will effectively decouple the emittance dilution from the transverse alignment of the quadrupoles and the BPMs. Thus, this technique, and in general, the idea of using the trajectory to cancel the emittance dilutions, will likely prove indispensable in future linear colliders.

References

[1] T. O. Raubenheimer and R. D. Ruth, "A Dispersion-Free Trajectory Correction Technique for Linear Colliders," submitted to Nucl. Instr. and Meth.; and SLAC–PUB–5222 (1990).

[2] V. Balakin, A. Novokhatsky, and V. Smirnov, Proc. 12th Int. Conf. on High Energy Accelerators, Fermilab (1983) p. 119.

[3] See any text on charged particle optics.

[4] The NLC is a 250 GeV on 250 GeV linear collider being studied at SLAC. Detailed parameters can be found in: Proc. DPF Summer Study, Snowmass '88 and Proc. Int. Workshop on Next–Generation Linear Colliders, SLAC–335 (1988).

[5] A. W. Chao, B. Richter, and C. Y. Yao, Nucl. Instr. and Meth. **178** (1980) 1.

[6] J. T. Seeman, "New Compensation of Transverse Wakefield Effects in a Linac by Displacing Accelerator Structures," presented at the 1990 Linear Accelerator Conf., Albuquerque, NM; and SLAC–PUB–5337 (1990).

[7] T. O. Raubenheimer and R. D. Ruth, work in progress.

[8] V. Balakin, private communication.

Table 1
Correction in the NLC

Method	ϵ_y	Trajectory rms
1-to-1	$22.9 \pm 21.3\ \epsilon_{y0}$	$72 \pm 3\,\mu$m
DF	$9.3 \pm 7.3\ \epsilon_{y0}$	$55 \pm 5\,\mu$m
WF	$1.09 \pm 0.05\ \epsilon_{y0}$	$44 \pm 3\,\mu$m

Figure Caption

1. The beam distribution after (a) 1-to-1, (b) DF, and (c) WF correction. The left-hand plots are the y-y' phase space while the right-hand plots are the beam in y-z space.

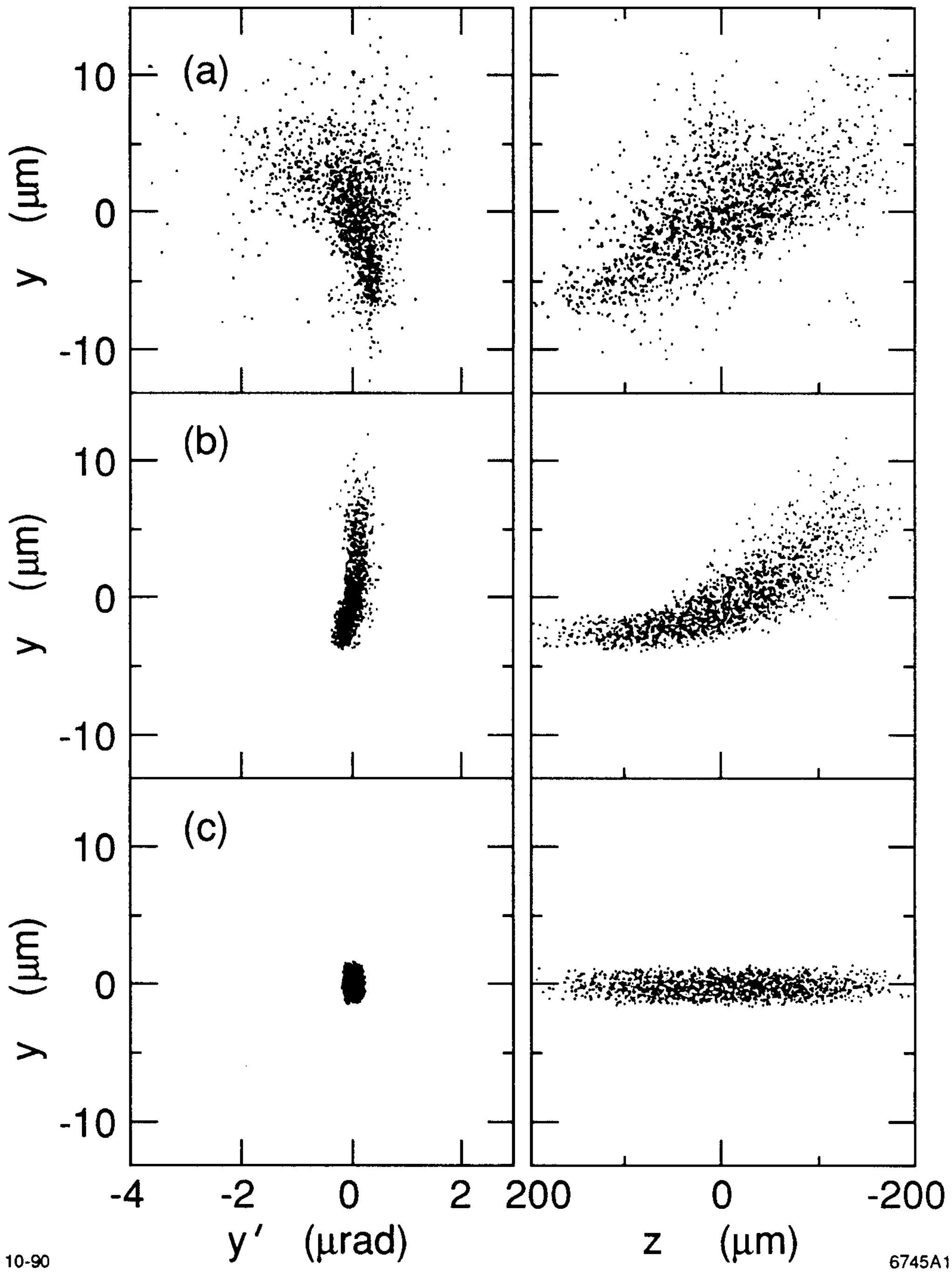

Fig. 1