Longitudinal crossover and the dynamics of uniform electron ellipsoids focused by a linear chirp

X. Xiang, P. M. Duxbury, and B. Zerbe

Longitudinal crossover and the dynamics of uniform electron ellipsoids focused by linear chirp

X. Xiang,* P.M. Duxbury, and B. Zerbe†

Department of Physics and Astronomy, Michigan State University

(Dated: December 22, 2020)

High resolution single-shot non-relativistic ultrafast electron microscopy(UEM) relies on adaptive optics to compress high intensity bunches using Radio Frequency (RF) cavities. We present a comprehensive discussion of the analytic approaches available to characterize bunch dynamics as an electron bunch goes through a longitudinal focal point after an RF cavity where space charge effects can be large. Methods drawn from the Coulomb explosion literature, the accelerator physics literature and the analytic Gaussian model developed for UEM are compared, utilized and extended in some cases. In particular the longitudinal focus may occur in two different regimes, a bounce-back regime and a crossover regime; and we characterize the critical point separating these regimes in the zero-emittance model. Results from N-particle simulations using efficient multipole methods are compared to the theoretical models revealing features requiring extensions of the analytic approaches; and in particular mechanisms for emittance growth and transfer are discussed.

I. INTRODUCTION

Modern ultrafast microscopy has the goal of resolving sub-picosecond time periods at sub-nanometer length scales [1, 2]. Consistently obtaining such resolution would allow scientists to visualize chemistry as it happens thus opening up a deeper understanding of mechanisms at the nano-scale that are important to life and modern technology[3, 4]. While a number of techniques are being explored to realize such microscopy [3–15], weakly-relativistic ultrafast electron microscopy (UEM), where the electron bunch has energies that are at most a significant fraction of the rest energy of the electron, has a number of attractive advantages. The first advantage is the engineering fact that the device needed for such experiments can be built on top of existing electron microscopes keeping additional engineering and expenses to a minimum. The second advantage is the physical fact that the use of strongly interacting electrons means that the number of electrons required to form an image is a relatively small number as compared to x-rays for example [12]. The ultimate goal is to reach the single-shot limit where the number of electrons in a bunch is large enough to form an image, but weakly-relativistic UEM also introduces technological hurdles as the space-charge effects of a high-density probing electron bunch is considerable at a number of points within the column[16–22]. These effects need to be characterized to provide an accurate model for design of high-intensity beamlines. In this paper we present such a model that can capture such space-charge effects at a longitudinal focal point.

While such so-called space-charge dominated regimes have been well described by accelerator physicists for cylindrical beams[23], the weakly-relativistic bunched nature, which can be thought of as ellipsoids with finite longitudinal extent, of the electrons in UEM requires additional tools. Some work has already characterized such dynamics in the non-relativistic regime near the electron source. Models of the longitudinal evolution of the bunch have been developed to describe the early dynamics of a bunch within an acceleration field before the center of mass motion becomes relativistic[24–28]; although we recently showed that the transverse dynamics should not be ignored when attempting to capture the important aspects of the electron bunch evolution[22]. Furthermore, once the bunch has expanded sufficiently, it is often argued that the internal space-charge effects become negligible; however, for weakly-relativistic UEM the bunch is recompressed resulting in the space-charge effects becoming significant at and near focal points where the density of the bunch is again high.

Fortunately, there have been tools developed in the astrophysics and Coulomb explosion literature where the mean-field effects of a uniform ellipsoidal electron bunch can be modeled through ordinary differential equations. Specifically, Lin et al. developed a model of gravitational collapse of an oblate ellipse that could be written as a system of differential equations for the ellipses' widths[29]. Similar techniques using the repulsive electrostatic force were developed by Grech et al. to model the inverse problem of Coulomb explosion[30]. Both techniques require a tractable force, and to simplify the analysis, both techniques assumed a uniform ellipsoid throughout the bunch evolution. However, such models assume a on-to-one relation between particle location and particle momentum, which is unphysical.

Separate from these efforts, Michalik and Sipe introduced an Analytic Gaussian (AG) model that is able to capture the effect of local momentum spread at every location [31–33]. To capture such effects, Michalik and Sipe utilize a measure common in accelerator physics, the normalized rms emittance, henceforth emittance, that represents the phase-space area occupied by the ensemble and which we will define mathematically later in the text. The AG model is presented in the reference

---

* xiangxuk@msu.edu
† zerbe@msu.edu

frame of the bunch, so it is only applicable as long as the bunch remains non-relativistic within the lab frame. Further, we argue here that the AG model is equivalent to the much older Kapchinskij-Vladimirskij (KV) envelope equations initially developed to describe the evolution of uniform ellipsoidal distributions[34]. Sacherer provided a simple perspective that showed that the KV envelope equations could be derived from basic, fundamental statistical considerations with applications of the mean-field force present from a uniform distribution[35], and the mathematical form of the AG model may be derived from similar considerations again assuming emittance conservation. We provide such a derivation later in this manuscript.

We have recently argued that emittance is an important measure in the statistical description of any ensemble of particles as it appears in the dynamics of the second order moments, as we will show here. We emphasize that the use of emittance is valid for understanding collective particle effects reflected in the statistical description of any physical situation — not just the one component plasma situation we investigate here. Of special note, the emittance is conserved for systems where all forces are perfectly linear, therefore ideal uniform distributions within the fields of gravitation and electromagnetism are of specific theoretical interest to the understanding of the dynamics of such ensembles[24]. Furthermore, the collective effects of an ensemble of particles may be conveniently partitioned into mean-field effects[36, 37], momentum spread effects[31, 35], finite size effects[38], and particle-particle scattering-like effects, and such effects drive changes within the emittance. Therefore, the analysis of emittance is fundamental in understanding the collective effects of any ensemble.

In this work, we have three primary goals: 1.) We extend the model of Grech et al. to capture focussing events, 2.) we place the envelope equations within the context of the UEM and accelerator physics literature, and 3.) we analyze changes in emittance during well-controlled focusing and crossover, as a leading indicator of typical non mean-field interaction effects within the ensemble. We start by extending the model employed by Grech et al. to include linear initial momentum-spatial correlations, aka chirp, in Sec.II. We show that the use of a chirp to introduce crossover can be treated precisely in the zero emmitance limit by an important extension of the mathematics utilized in the paper of Grech et al. We call this the modified Coulomb explosion (MCE) model in the text. We find that this MCE model naturally leads to the concept of a critical chirp that describes a collective behavior transition for particles within this model. Next, in Sec.III, we derive the AG formalism from a statisti-

cal vantage point assuming a linear force. We explicitly demonstrate how the Gaussian assumption differs from the uniform assumption only by a constant that can be absorbed into the number of particles in the bunch if the model is used to represent experimental data for example (see Appendix B 2). Further we point out that the envelope equations we derive from this statistical perspective are a generalization of the MCE model. This observation allows us to partially disentangle the effects of the collective self force and momentum spread on the predicted dynamics of the bunch, and we analyze some important physics of bunch evolution using this insight. As the deviation of real systems from such a model is due to stochastic effects that simultaneously result in emittance change, the theoretical predictions are compared to N-particle simulations where such stochastic scattering events are present. Consistent with previous theory[23], we show that emittance is transferred from hotter to colder dimensions; however, we also show that emittance increases almost simultaneously in both the transverse and longitudinal directions around crossover when the initial chirp is larger than the critical chirp. We note that this can not be explained through the standard mechanism of heat transfer; and we postulate two mechanisms that may be important in the collective dynamics of a focussed, charged bunch.

II. SPATIAL EVOLUTION

We revisit Grech et al.'s model for Coulomb explosion. Broadly, this model assumes that the force acting on a particle within the uniform ellipsoidal ensemble is the mean-field force calculated by the application of Laplace equations to a uniform distribution of electrons. The modification we introduce is an initial linear relationship between the initial position and the initial velocity of the particle, which we will call the "chirp", and this modification naturally leads to the identification of a critical chirp that demarcates two qualitatively different regimes of bunch behavior within this model.

A. The mean-field framework

We first recall the well-known quadratic form of the electrostatic potential for position (x, y, z) inside a uniform electron ellipsoidal bunch with semi-axes of (a, b, c) and charge number density n that can be obtained using Laplace's equations[39]:

$$V(x, y, z) = \frac{n \cdot abc \cdot e}{4\varepsilon_0} \cdot \int_0^\infty \left(1 - \frac{x^2}{a^2 + s} - \frac{y^2}{b^2 + s} - \frac{z^2}{c^2 + s}\right) \frac{ds}{\sqrt{(a^2 + s)(b^2 + s)(c^2 + s)}}, \tag{1}$$

where ε_0 is the vacuum permittivity. We assume rotational symmetry about the z axis enabling us to introduce the radial coordinate $r = \sqrt{x^2 + y^2}$. Although the detailed calculations below are performed specifically for prolate ellipsoids ($a = b < c$), similar results are valid for general uniformly charged ellipsoidal bunches.

The electrostatic field may be obtained from Eq. (1) using $\vec{E} = -\vec{\nabla}V$. Due to the symmetry, the angular portion of the field is 0. Thus the electric field may be written as:

$$\vec{E}(r,z) = E_r(r)\hat{r} + E_z(z)\hat{z} \qquad (2)$$

with $\hat{r}$ and $\hat{z}$ representing the radial and longitudinal unit vectors, respectively, and

$$E_r(r) = \frac{ne}{2\varepsilon_0}\xi_r(\alpha) \cdot r \qquad (3a)$$

$$E_z(z) = \frac{ne}{2\varepsilon_0}\xi_z(\alpha) \cdot z \qquad (3b)$$

where $\alpha = a/c$ is the ellipsoid aspect ratio and the corresponding geometry coefficients $\xi_r(\alpha)$ and $\xi_z(\alpha)$ are

$$\xi_r(\alpha) = \alpha^2 \int_0^\infty \frac{ds}{(\alpha^2 + s)^2(1+s)^{1/2}} \qquad (4a)$$

$$\xi_z(\alpha) = \alpha^2 \int_0^\infty \frac{ds}{(\alpha^2 + s)(1+s)^{3/2}} \qquad (4b)$$

Provided that the initial velocity of the particles in the ellipsoid can be expressed as linear functions of their coordinates, the linear relation between the electric field felt by a particle and the particle's position results in the preservation of the uniformity of the ellipsoidal bunch. This greatly simplifies our analysis as the formulation presented above applies to the bunch for all time and the evolution reduces to the determination of two degrees of freedom. Specifically, the temporal evolution of the entire bunch can be represented by the evolution of two dimensionless scaling functions, $R(t)$ and $Z(t)$, i.e. the trajectory of any particle with initial position (r_0, z_0) inside the uniform ellipsoid is given by $(r_0 R(t), z_0 Z(t))$, where R and Z are independent of the initial position (r_0, z_0). Thus, the parameters for describing the bunch changes accordingly: (i) the semi-axis of the ellipsoids can be written as $(a, c) = (a_0 R, c_0 Z)$, (ii) the transient aspect ratio can be written as $\alpha(t) = \alpha_0 \cdot R/Z$, (iii) the number density can be derived using conservation of charge $N_{total} = n_0 \cdot (4\pi/3)a_0^2 c_0 = n(t) \cdot (4\pi/3)a^2 c$ giving $n(t) = n_0/(R^2 Z)$, and (iv) the spatial variance of the bunch change to $\sigma_z^2(t) = \sigma_{z0}^2 \cdot Z^2$ and $\sigma_r^2(t) = \sigma_{r0}^2 \cdot R^2$. Therefore, it should be apparent that any parameter in the problem can be determined from R and Z, which we set out to determine for all time.

In the non-relativistic limit, the equations of motion (EOM) of a particle inside the field determined in Eq. (3) can be simply determined using $\ddot{\vec{x}} = \frac{q}{m}\vec{E}$. These EOM reduce to two dimensionless ordinary differential equations (ODEs) for our scaling parameters:

$$\frac{d^2 R}{d\tau^2} = \frac{\xi_r(\alpha)}{RZ} \qquad (5a)$$

$$\frac{d^2 Z}{d\tau^2} = \frac{\xi_z(\alpha)}{R^2} \qquad (5b)$$

with dimensionless reduced time,

$$\tau = t \cdot \sqrt{\frac{e^2 n_0}{2\varepsilon_0 m}} = t \cdot \Omega_0 \qquad (6)$$

and electron mass m. Notice that: (i) the time scaling factor $\Omega_0 = \frac{1}{\sqrt{2}}\omega_{p0}$ where $\omega_{p0}(n_0) = \sqrt{\frac{e^2 n_0}{\varepsilon_0 m}}$ is the initial plasma frequency and (ii) the geometry coefficients ξ_r and ξ_z solely depend on the aspect ratio α rather than specific value of a and c. This means that starting with the same initial conditions for the ODEs, bunches with the same initial aspect ratio α_0 but different initial density n_0 will lead to identical behaviors only differing by the time scaling factor Ω_0 determined by the initial number density n_0. Eq. (5) are more or less the ODE's used by Lin et al.[29] and Grech et al.[30] except we have scaled the time to be more general, so the model we have presented so far does not significantly differ from those works.

B. Initial conditions

The behavior predicted by a specified system of non-chaotic ODE's is entirely determined by its initial conditions, and the initial conditions we consider are

$$R(\tau = 0) = 1 \qquad (7a)$$

$$Z(\tau = 0) = 1 \qquad (7b)$$

$$\left.\frac{dR}{d\tau}\right|_{\tau=0} = -\nu_r^* \qquad (7c)$$

$$\left.\frac{dZ}{d\tau}\right|_{\tau=0} = -\nu_z^* \qquad (7d)$$

where ν_i^* is trivially proportional to the linear chirp. We call ν_z^* the reduced longitudinal chirp, and its proportionality to the linear chirp can be obtained by noting $p_z(z_0) = m \cdot (\nu_z^* \Omega_0 \cdot z_0) = m\mathcal{C}_z \cdot z_0$ where $\mathcal{C}_z$ is the longitudinal linear chirp. Notice that Eqs. (7a) and (7b) represent the initial scaling of the ellipsoid and are by definition set to 1 as these parameters represent the scaling of the transverse and longitudinal dimensions, respectively, from their initial values. On the other hand, Eqs. (7c) and (7d) represent the initial rate of change of the scale functions R and Z, which can be roughly thought of as the velocity of the expansion. Lin et al. and Grech et al. set $\nu_r^* = \nu_z^* = 0$ to model gravitation collapse and Coulomb explosion, respectively, where the bunch is assumed to start from rest. The Coulomb explosion results

were found to be in good agreement with molecular dynamics (MD) simulations for time-dependent energy distributions and particle-in-cell (PIC) simulations for temporal ellipsoid radii evolution[30]. For our purposes, we assume $\nu_r^* = 0$ and $\nu_z^* \neq 0$ to model the effect of a longitudinal lens, e.g. a RF cavity. Specifically, notice that if the reduced longitudinal chirp is positive, i.e. $\frac{dZ}{d\tau}\big|_{\tau=0}$ is negative, Z will initially decrease, and the bunch will be focused in the longitudinal direction. In summary, the focusing process of a uniform charged ellipsoid is entirely determined by its initial aspect ratio and its reduced longitudinal chirp as the density of initial bunch determines only the time scale of the evolution. We call this general form of Grech et al.'s model the modified Coulomb explosion (MCE) model.

In particular note that the reduced longitudinal chirp is dimensionless while the longitudinal chirp has dimensions of inverse time. This is because the reduced chirp is the actual chirp scaled by Ω_0, and this cancels the time dimension. As Ω_0 depends solely on density, the reduced chirp is more general as the density determines the time scale and therefore the ODE represents the interplay between the geometry and the electrostatic force. However, if the density is not important in our discussion of some physical observation, we will often drop the "reduced" when discussing the chirp as the statement should apply to both the reduced chirp as well as the actual chirp.

C. Critical reduced chirp

As the effect of aspect ratio on the evolution has been well studied previously[30], we examine the effect of the reduced longitudinal chirp on the bunch focusing of a pre-specified aspect ratio, $\alpha_0 = 10/75$. Specifically, we are interested in modeling the bunch reaching a minimum in longitudinal extent after $\tau = 0$ which occurs when $-\nu_z^* < 0$. We define the time to focus, τ_f, as the dimensionless time at which the bunch reaches its minimum longitudinal width. As can be seen in Fig. 1, τ_f is a function of the reduced chirp, i.e. $\tau_f = \tau_f(\nu_z^*)$.

Furthermore, define Z_f to be the longitudinal scaling parameter at the focal point, i.e. $Z_f = Z(\tau_f) = min(Z(t))$. Moreover notice that $Z(t) \geq 0$, so $Z_f \geq 0$; in fact, for sufficiently large reduced chirps $Z_f = 0$ as can be seen in Fig. 1. This is because the evolution of the longitudinal scaling parameter, seen in Eq. (5b), is dependent only on $1/R^2$, and $R > 1$ in our model. This means that if ν_z^* is sufficiently large, the initial longitudinal chirp overcomes the repulsion of the electrostatic force and the bunch briefly collapses to a two-dimensional object at the focal point. We call the smallest magnitude of the reduced chirp for satisfying this condition the critical chirp, $\nu_{z,c}^*$. Notice that $\tau_f(\nu_{z,c}^*) = \infty$ and that $\frac{d\tau_f}{d\nu_z^*} \begin{cases} > 0, & \nu_z^* < \nu_{z,c}^* \\ < 0, & \nu_z^* > \nu_{z,c}^* \end{cases}$ again as can be seen in Fig. 1.
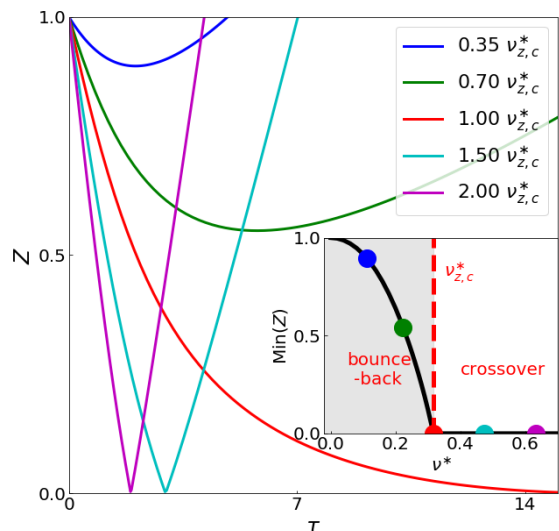

FIG. 1: Longitudinal width evolution $Z = Z(\tau)$ of prolate ellipsoids with $(\alpha_0 = 10/75)$ driven by different initial chirps in numeric solutions of the MCE model ranging from below the critical chirp $(0.35\ \nu_{z,c}^*)$ to well above $(2.0\ \nu_{z,c}^*)$. The sub-graph shows the dependence of minimum width on initial reduced chirp. The red dot represents the critical value $\nu_{z,c}^*$ for the particular $\alpha_0 = 10/75$, and the bounce-back and crossover regimes a separated by the vertical, dashed, red line at this point.

In other words, the behavior of the model can be partitioned into two categories characterized by whether the initial longitudinal chirp is greater than or less than the critical chirp. More specifically, as the magnitude of focusing chirp is increased from 0, the minimum width of the bunch decreases and the time to focus increases. This trend continues until the critical chirp is reached where the corresponding time to the focal point becomes infinitely large, i.e. $\tau_f \to \infty$ as $\nu_z^* \to \nu_{z,c}^*$. Above the critical chirp, the bunch will overcome the Coulomb repulsion and be compressed through a longitudinal crossover as electrons starting from one side of the bunch cross the center of mass and then begin to expand on the other side. We refer to this as the "crossover" regime, and in this regime further increasing the chirp has no effect on the 0 minimum width but decreases the time to focus. In contrast, we call the regime below the critical chirp the "bounce-back" regime as a particle within the bunch with such a chirp follows a trajectory that reverse its initial direction.

The crossover event adds complexity to simulations of the model. Specifically, the linearity of both the force and the velocities of the particles in the model indicates that all the crossover incidents happen simultaneously across the bunch at τ_f within the crossover regime, creating a 2D singularity in the EOM with $Z \to 0$. Before the crossover, the chirp is negative while after the crossover the chirp becomes positive. As the force in the

z-direction is very small due to geometric considerations, the speed of the particles do not change substantially, just the sign of the linear relationship in phase space. This necessitates careful treatment of the chirp through the crossover event. We accomplish this treatment by using a small time step to propagate the EOM up until Z goes below zero. As Z is a scale, the negative sign has no physical meaning and indicates that crossover occurred within the previous time step. So, we stop the simulation and flip the value of both longitudinal position scaling, Z, and longitudinal momentum, p_z. After this, the same EOM are used to integrate the parameters. In effect, this skips the singularity by an infinitesimal step size in time. In addition, this also implies that the crossover case, where Z will pass through 0 in this fashion, will have a sudden change in longitudinal chirp as compared to the bounce-back case where such Z does not pass through 0 and the chirp instead smoothly changes due to the effect of the repulsive mean-field force.

Analogous to our longitudinal treatment, a radial chirp can also be added by setting ν_r^* in Eq.7c to a non-zero value. Furthermore, this treatment may be combined with the longitudinal chirp to model bunches focused in both degrees of freedom simultaneously – a treatment that is outside of the scope of this paper. However, in contrast to the longitudinal dimension, this model predicts that there is no such critical chirp or crossover in radial focusing. This occurs because of $d^2 R/d\tau \propto 1/R$ as can be seen in Eq. (5). This indicates that the force in the transverse direction diverges as the bunch focuses radially preventing the singularity in the longitudinal scaling parameters seen when only the longitudinal direction is focused.

In the general situation, this difference between being able to focus through a singularity longitudinally but not transversely is a result of attempting to focus two dimensions, i.e. $\hat{x}$ and $\hat{y}$, simultaneously. Focusing in more than one dimension in this model is not possible even when all dimensions are treated separately as the Coulomb repulsion on one dimension is inversely dependent on the widths of the other two dimensions. In other words, there is only the bounce-back regime when more than one dimension is focussed concurrently. We will later (in Sec.III) discuss how emittance influences the minimum width of the bunch, and this statistical measure reintroduces the ability of particles to crossover even when the bunch is in the bounce-back regime. For the rest of this manuscript, though, we will only focus on the longitudinal focusing where both the crossover and bounce-back regimes are accessible in the model.

One important feature of the critical reduced chirp, $\nu_{z,c}^*$, is its exclusive dependence on the initial aspect ratio α_0. This fact stems from the governing EOM solely depending on the aspect ratio. In Fig. 2, we present the reduced critical chirp as a function of the initial aspect ratio. Specifically, note that for large α_0 often referred to as the "pancake" regime [22, 24–28], we always have $\alpha \gg 1$, where the geometry coefficients ξ_r and ξ_z can be

approximated in closed forms, with

$$\xi_r(\alpha \to \infty) \simeq \pi/(2\alpha) \to 0 \tag{8a}$$

$$\xi_z(\alpha \to \infty) \simeq 2 - \pi\alpha^2(\alpha^2-1)^{-3/2} \to 2 \tag{8b}$$

Therefore, the longitudinal motion can be treated as the elementary constant acceleration kinematic equation:

$$Z(\tau) = Z_0 - \nu_{z,c}^*\tau + \frac{1}{2} \cdot \xi_z(\infty) \cdot \tau^2 \tag{9}$$

with $\tau_f = \nu_{z,c}^*/\xi_z(\infty)$ and $Z(\tau_f) = 0$. In essence, this equation corresponds to the longitudinal crossover within the planar model, which gives an analytic critical reduced chirp of 2. This value corresponds to the asymptote seen Fig. 2. That is, for sufficiently large aspect ratios, the envelope model may be analyzed for small amounts of time using the planar model, which should be unsurprising.
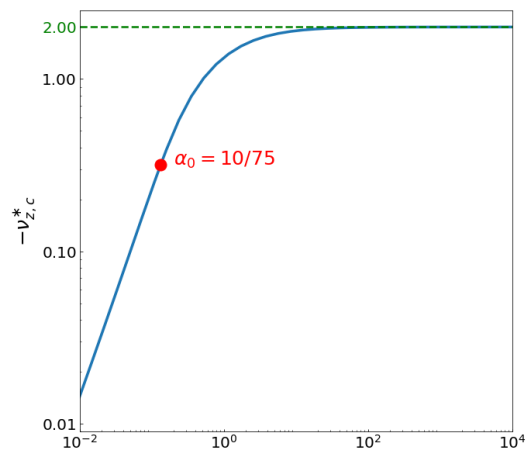

FIG. 2: Dependence of critical reduced longitudinal chirp, $\nu_{z,c}^*$, on the initial aspect ratio, α_0. The green dashed line represents the horizontal asymptote, $\eta_{z,c}^*(\alpha_0 \to \infty) = 2$. The red dot represents the aspect ratio of our MD simulation.

III. ENVELOPE EQUATIONS

In this section, we present a brief derivation of the envelope equations, and we compare this model to the Analytic Gaussian (AG) formalism. Our derivation is essentially identical to Sacherer's derivation of the KV envelope equations[35] and closely follows our recent derivation presented with our discussion of the sample perspective[38].

A. Derivation

We first introduce the statistics of the bunch and their dynamics. In each degree of freedom (x, y, z), we need three quantities to describe the second order statistics of

the phase space: s_i, s_{p_i} and s_{i,p_i}, with $i = T, z$ for the transverse ($T = x, y$) or longitudinal direction. The basic statistics are then

$$s_i^2 = \overline{i^2} - \bar{i}^2 \tag{10a}$$

$$s_{p_i}^2 = \overline{p_i^2} - \bar{p}_i^2 \tag{10b}$$

$$s_{i,p_i} = \overline{ip_i} - \bar{i}\bar{p}_i \tag{10c}$$

where the bar operator indicates the mean, e.g. $\overline{xp_x} = \frac{1}{N}\sum_{j=1}^{N} x_j p_{x,j}$. As the number of particles is a constant, derivatives commute with sums, and derivatives of products can be determined by the chain rule, we have $\frac{d}{dt}\bar{a} = \overline{\frac{da}{dt}}$ and it is straightforward to show

$$\frac{d}{dt} s_{a,b} = s_{\frac{da}{dt},b} + s_{a,\frac{db}{dt}} \tag{11}$$

Thus the time derivatives of our phase-space statistics are

$$\frac{ds_i^2}{dt} = \frac{2}{m} s_{i,p_i} \tag{12a}$$

$$\frac{ds_{i,pi}}{dt} = s_{i,F_i} + \frac{1}{m} s_{p_i}^2 \tag{12b}$$

$$\frac{ds_{pi}^2}{dt} = 2s_{p_i,F_i} \tag{12c}$$

assuming non-relativistic dynamics.

We have been calling this system of equations the second order statistical kinematics as they exactly specify the evolution of the second order statistics in the same way as kinematics exactly describes the evolution of a single particle. Notice that each dimension has 3 degrees of freedom — these degrees of freedom describe the $2D$ covariance matrix in (x,p_x) space, etc. Moreover, alternative choices of parameters, other than the covariance matrix elements above, may be chosen. In fact, one such well known choice of these degrees of freedom are the Courant-Snyder parameters used in accelerator physics; we provide context on why we chose not to use the Courant-Snyder parameters for this manuscript and how the statistical kinematics inform the evolution of the Courant-Snyder parameters including the effect of adiabatic damping in Appendix A. However, the statistical kinematics apply generally and are important in the understanding of non-equilibrium dynamics in any field; in this manuscript, we look at parameters that are more in line with existing literature in the UEM community.

Specifically, we compare the second order statistical kinematics to the three degrees of freedom represented in the AG model. In Appendix B 1, we show that the AG model can be simply derived from the statistical kinematics without integration of the phase space. Specifically, a single assumption reproduces Michalik and Sipe's published model from the statistical kinematics; namely, assume that the force on a particle in the AG model is linear and can be written as

$$F_i(i) = \frac{1}{4\pi\epsilon_0} \frac{Ne^2}{6\sqrt{\pi}\sigma_i^3} L_i(\xi)i. \tag{13}$$

Further details showing this explicitly can be found in Appendix B 1.

However, this linear force assumption has 2 somewhat subtle, and related, problems. The first has to do with the description of the relationship between the position of a particle in the distribution and the force it experiences on average. The line of best fit has slope $\frac{s_{i,F_i}}{s_i^2}$, so s_{i,F_i} can be thought of as the slope of the best fit line times the spatial variance. While the slope of the line for a Gaussian is essentially described by the force in Michalik and Sipe's AG model, we note that the assumption that the distribution will remain Gaussian has been found to be incorrect[22]. Specifically, the slope of the best fit line is specific to the given distribution, and as the Gaussian distribution evolves toward a ringed distribution, the slope of this line changes partially just in response to this change in distribution. This issue can be partially avoided theoretically by assuming a uniform distribution that does continue to be uniform as it evolves — at least in the continuum, mean-field, non-relativistic, zero-emittance limit. For this work, we assume $F_i(i) = (m\omega_{p_0}^2/2)\xi_i(\alpha)i$ which leads to the envelope equations we use as well as the equations used by Sacherer[35]. The difference between the AG model and these uniform envelope equations can be mathematically shown to be nothing more than a difference in a dimensionless constant relating the force used in the analysis of Michalik and Sipe, the force we use. In Appendix B 2, we calculate this constant and find that it is only 1.05 indicating that these models are essentially equivalent. Because these two models differ by only about 5% in the force, either set of equations can be used in most applications to experiment.

However, the linear assumption results in a more serious issue. Specifically the force in a real bunch differs from the linear approximation. This is even true for the uniform distribution although the more consistent non-linearities of the Gaussian distribution result in more significant deviations. So while the s_{x,F_i} can be captured in many situations with the mean-field force, it is important to understand these deviations especially for the term relating the momentum and the force, $\frac{m^2 c^2}{s_i^2}\frac{d\varepsilon_{i,p_i}^2}{dt} = \frac{2}{s_i^2}\left(s_i^2 s_{p_i,F_i} - s_{i,p_i} s_{i,F_i}\right)$, cannot be likewise captured. As the Gaussian distribution results in significant deviations between the linear force and the mean-field force, the rms emittance has additional effects on the evolution of the emittance growth than the uniform distribution's purely stochastic driven emittance changes. This second point has not been examined in the literature, and we begin to evaluate the stochastic driven aspects of emittance growth in this manuscript by investigating the emittance change during simulation as compared to the uniform envelope equation predictions. As the gold standard for simulating realistic forces is N-particle simulation, we conduct N-particle simulations in our exploration of these effects.

In terms of understanding the physics, it is convenient

to introduce a variable representing the average local variance in the momentum,

$$\eta_i = \sqrt{s_{p_i}^2 - \frac{s_{i,p_i}^2}{s_i^2}} \tag{14}$$

Notice that η_i^2 has dimensions of momentum and $\frac{1}{2m}\eta_i^2$ has dimensions of energy. Furthermore, the emittance can be written as $\epsilon_{i,p_i} = \eta_i s_i$ similar to Michalik and Sipe's notation[31]. With this notation and an assumption of a linear force, $F_i = \mathcal{K}_{Fi}(\alpha) \cdot i = (m\omega_p^2/2)\xi_i(\alpha)i$, our system of ODEs becomes

$$\frac{ds_i^2}{dt} = \frac{2}{m} s_{i,p_i} \tag{15a}$$

$$\frac{ds_{i,p_i}}{dt} = \frac{1}{m}\left(\frac{s_{i,p_i}^2}{s_i^2} + \eta_i^2\right) + \mathcal{K}_{Fi}(\alpha)s_i^2 \tag{15b}$$

$$\frac{d\eta_i^2}{dt} = -\frac{2s_{i,p_i}\eta_i^2}{ms_i^2} \tag{15c}$$

Although the above derivations have been performed for the Coulomb interaction, we would like to stress that the same conclusions can be drawn for any interaction that leads to linear dependence between force and position. Additionally, the generalization to any general ellipsoid is simple using three degrees of freedom with $i = X, Y, Z$ and corresponding geometry coefficients (ξ_x, ξ_y, ξ_z) as functions of the ratio between three axes.

B. Dynamics of the envelope equations

Notice that the non-interacting bunch model can be obtained by setting $\mathcal{K}_{Fi}(\alpha) = 0$ in Eq. (B2). Further, notice that the MCE model gives identical predictions to the zero-emittance limit of the envelope equations above, as can be seen in in Fig.3.

The bunch dynamics can be better understood analytically by investigating the dynamics of the linear chirp, $\mathcal{C}_i = \frac{s_{i,p_i}}{s_i^2}$,

$$\frac{d}{dt}\mathcal{C}_i = \mathcal{K}_{Fi}(\alpha) + \frac{mc^2\varepsilon_{i,p_i}^2}{(s_i^2)^2} - \frac{1}{m}C_i^2 \tag{16}$$

That is, the chirp evolution is influenced by three factors. These factors can be decomposed into an interacting effect, $\mathcal{K}_{Fi}(\alpha)$, and two expansion effects: (a) expansion due to velocity spread, $\frac{mc^2\varepsilon_{i,p_i}^2}{(s_i^2)^2}$, and expansion due to the chirp, $-\frac{1}{m}C_i^2$. This realization leads us to identify four ways to investigate the physics of the bunch dynamics: 1.) the non-interacting model with 0 emittance where the dynamics are entirely driven by the chirp, 2.) the non-interacting model with non-zero emittance where the dynamics are driven by existing emittance and the chirp, 3.) the MCE model or equivalently the envelope equations with 0 emittance whose bunch dynamics are
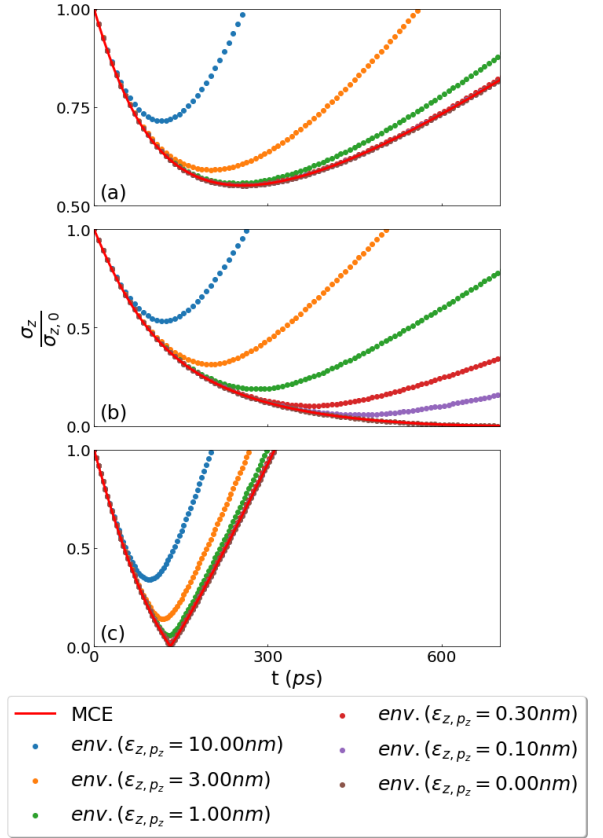

FIG. 3: Longitudinal width evolution, σ_z, divided by the initial longitudinal width, $\sigma_{z,0}$, of prolate ellipsoids with $(\alpha_0 = 10/75)$ focused by different initial chirps: (a) $0.7\ \nu_{z,c}^*$, (b)$1.0\ \nu_{z,c}^*$, (c) $1.5\ \nu_{z,c}^*$. In each figure, the red solid line represents the prediction from the MCE model and the dotted lines represent the envelope equations with different emittance ranging from 0 to 10 nm. Notice that the envelope equations with zero emittance and the MCE model are in perfect agreement. Also notice that (1) increasing the emittance makes the waist larger and moves it earlier and (2) the evolution of the width statistic is more responsive to emittance when the chirp is in the vicinity of the critical chirp.

driven entirely by the internal Coulomb repulsion and existing chirp, and 4.) the full envelope equations where all three effects and their interaction effects may be examined. We discuss the effects of these contributions in the rest of this section by isolating them.

Notice that in the case that all other terms except the chirp term of Eq. (16) are zero, then we can obtain an analytic solution for the chirp:

$$C_i = \frac{1}{C_{i,0}t - 1}C_{i,0} \tag{17}$$

where $C_{i,0}$ represents the initial chirp in the ith direction at time 0. Notice that there is a singularity for this solution when $t = \frac{1}{C_{i,0}}$. Before this singularity, the bunch

is moving in and the chirp is becoming more negative. After this singularity, the chirp is positive, the bunch is moving out, and the chirp approaches 0.

Notice that the emittance term is always positive, that is it is driving the chirp to increase in time. In isolation, this is because local velocity spread, for which $\frac{c^2 \varepsilon_{i,p_i}^2}{s_i^2}$ is an estimate, drives faster particles to move out more quickly thus resulting in a higher correlation between the spatial and momentum components. If emittance is conserved and present along with the chirp term, the evolution of chirp will never pass through the singularity as the velocity spread will increase inversely proportionally to the inverse quartic power of the size of the bunch and will eventually overwhelm the squared chirp contribution. At the point where they are equal, i.e. $C_i^2 = \frac{m^2 c^2 \varepsilon_{i,p_i}^2}{(s_i^2)^2}$, the chirp turns around and begins to either increase (if chirp is negative) or decrease (when chirp is positive). It may be argued that the region between the two points where the emittance and chirp are equal can be thought of as the emittance dominated portion of the dynamics.

The force term in isolation is generally understood. The outward repulsion of the Coulomb force drives a chirp increase. This repulsion decreases as the bunch expands and thus has a similar interaction with the chirp as the emittance, which is likewise strictly positive and decreases with expansion; of course, the details of where this effect dominates differ as the effect on the chirp falls roughly, with a geometric effect, inversely to the cubic power of the size of the bunch. Thus the emittance and the electric repulsion cooperate to increase the chirp and expand the bunch. If we were to consider gravitation instead, the sign of the force effect flips and the emittance and force effect react antagonistically.

In Fig. 3 we vary the emittance in the envelope model to specifically demonstrate its role on the dynamics. We can see that the width trajectories largely follow one another before the larger emittance predictions break off and reach their minima at a short time later. This is due to the fact that the emittance term increases rapidly with decreases in the bunch size. While there is a small effect of this emittance effect on the bunch width, to first order, it is primarily the shift in the time to the focus that increases the size of the waist.

The kinetic energy can also be modeled through the envelope equations. The kinetic energy can be exactly written as $KE = \sum_i \frac{p_i^2}{2m}$ and can be decomposed into $KE = KE_x + KE_y + KE_z$ where $KE_i = \frac{N}{2m}(p_{i,CoM}^2 + s_{p_i}^2)$ and where $p_{i,CoM}$ is the momentum of a single particle at the center of mass of the bunch in the ith direction. Assuming the center of mass momentum doesn't change, the kinetic energy evolution along the ith dimension can be written as:

$$\frac{d}{dt}KE_i = \frac{N}{2m}\frac{ds_{p_i}^2}{dt}$$
$$= \frac{N}{m}s_{i,p_i}\mathcal{K}_{Fi}(\alpha), \qquad (18)$$

in any linear model. That is, the kinetic energy can be transferred via the mean-field force into or out of the potential. Furthermore, as the different components (x, y, and z) can be independently controlled so that in effect energy is being transferred into the potential by one component but out of the potential by another, this mechanism can lead to kinetic energy width transfers between the dimensions. The result of this mechanism can be seen in Fig. 4 for a bunch in the crossover regime. Fur-
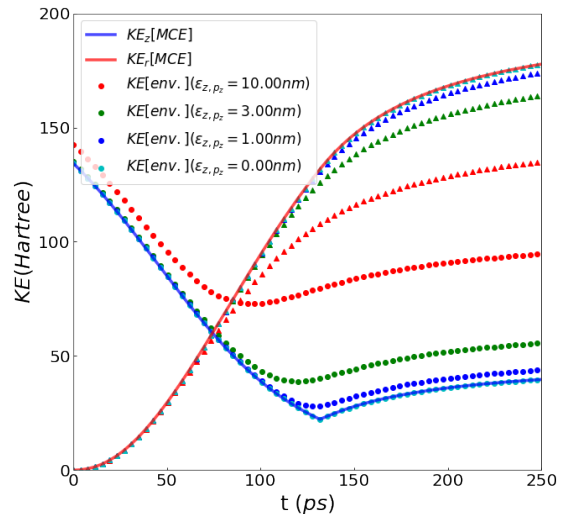

FIG. 4: The longitudinal and transverse kinetic energy for the crossover case ($1.5 \nu_{z,c}^*$ corresponding to panel (c) in Fig.3), with solid lines for the MCE model and dotted lines for the envelope equations with different emittance (circle for KE_z and triangle for KE_T). The sudden change of direction for the slope of the MCE model prediction of the longitudinal kinetic energy comes from the sign flip of chirp discussed in Sec.II.

thermore, notice that the effect of the emittance can also be seen Fig. 4. Specifically, the transfer of kinetic energy between the components is reduced by increasing the emittance. This occurs as the non-zero emittance results in the minimum width being both larger and occurring earlier than the MCE model. In turn, this larger size reduces the forces experienced reducing this transfer. Moreover, the earlier focal time results in the transfer finishing earlier, and all of these effects are factors in determining the amount of energy transferred between dimensions. We will provide more details of this aspect of the model in future publications.

IV. N-PARTICLE SIMULATIONS

Now that details of the models are understood, we compare the models to N-particle simulations. While the models describe the evolution of the bunch under specific conditions, i.e. conserved emittance for the envelope equations and zero emittance for the MCE model, no

such assumption is present in the N-particle simulations. The only assumption that we make in these simulations is that the bunch remains non-relativistic and thus electrostatics can be used to model the inter-particle interaction.

The simulations were conducted using in-house code. This code has been validated through comparison to other in-house code implementing the PIC algorithm from Warp[22]. This code employs the non-relativistic equations of motion for every electron using velocity Verlet integration where we used the Fast Multipole Method (FMM) from the fmmlib3d library[40] to calculate the field. As emittance increases initially due to disorder-induced heating (DIH) [41], the bunch needs to equilibrate before the focusing simulation. We first place electrons inside a simulation box with periodic boundaries at the target density, which is 10,000 electrons for a prolate ellipsoid with the semi-axes of $(10\mu m, 10\mu m, 75\mu m)$. The starting position of the electrons are randomly drawn from a uniform distribution and the starting momentum is zero. Then the electrons are thermalized using Particle-Particle-Particle-Mesh (PPPM) methods in LAMMPS[42] (http://lammps.sandia.gov) for over 10 plasma oscillation periods. At the end of thermalization, we select electrons inside the desired prolate ellipsoidal region to construct one sample of initial conditions. To mitigate the stochastic effects in this procedure, we prepare 90 such samples. This process results in 90 ellipsoids of particles with non-zero emittance that experience only minor additional DIH at the beginning of the focusing simulation. We call these initial conditions "warm" due to the non-zero emittance.

A. Longitudinal width and kinetic energy evolution

Simulations were performed with 3 representative initial chirps: (i.) $0.7\ \nu_{z,c}^*$ in the bounce-back regime, (ii.) $1.0\ \nu_{z,c}^*$ at the critical chirp, and $1.5\ \nu_{z,c}^*$ within the crossover regime. N-particle simulations were conducted by first thermalizing the bunch without chirp and with periodic boundary conditions before allowing the bunch to evolve with the appropriate chirp. The average initial phase-space statistics of the bunch post-thermalization were used to initialize the envelope equations.

The envelope model prediction is largely in agreement with the N-particle simulation except for the simulation at the critical chirp as can be seen in Fig. 5, and this prediction deviates most significantly at the focal point. Similar results can be seen with the kinetic energy evolution except for some noticeable deviation in the prediction for the bounce-back regime $(0.7\ \nu_{z,c}^*)$. These discrepancies arise from momentum spread as supported by re-examining the envelope equations using the average emittance from the N-particle simulation at every time step (dotted lines).

Further as shown in Fig. 5, the envelope equation predictions deviate from the N-particle simulations in three
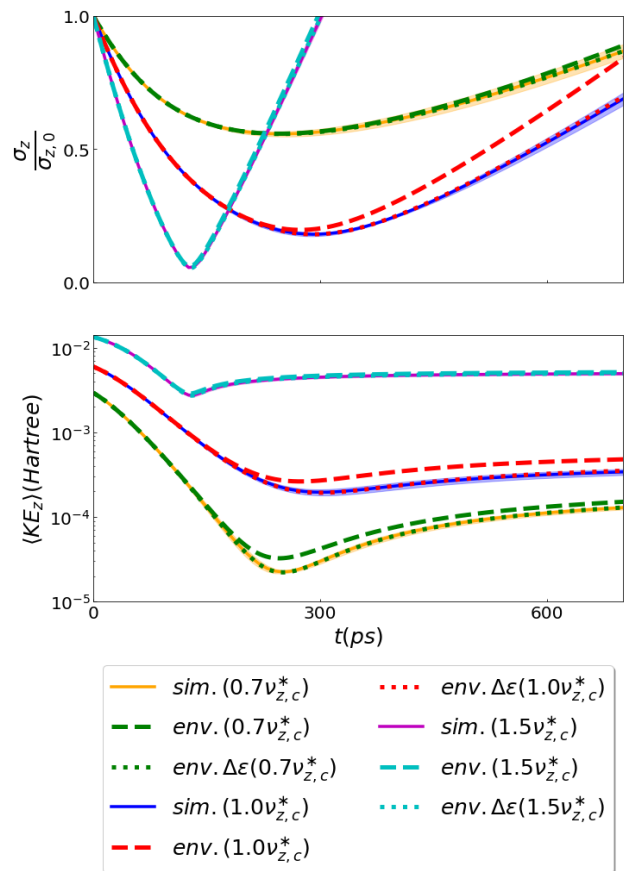

FIG. 5: Comparison of spatial width evolution and average kinetic energy in the longitudinal direction, $\frac{1}{N}KE_z = \frac{1}{2m}\left(\eta_z^2 + \frac{s_{z,p_z}^2}{s_z^2}\right) = \frac{1}{2m}s_{p_z}^2$, of the bunch focused by three different initial chirps, $0.7\ \nu_{z,c}^*$, $1\ \nu_{z,c}^*$, and $1.5\ \nu_{z,c}^*$. The line style of the plot indicates the simulations type: solid = mean of 90 N-particle simulations with the region shaded within 1 standard deviation of the mean (sim), dashed = envelope equation with conserved emittance (env), and dotted = envelope equation with emittance provided from simulation (env $\Delta\varepsilon$).

aspects: 1.) a slightly larger minimum width occurring at 2.) an earlier t_f with 3.) a faster expansion after the focal point. These deviations are most significant at the critical chirp, where the bunch evolution is most sensitive to changes in velocity spread measured by emittance as we discussed in a previous section. We previously saw similar trends in the minimum width, the time to the minimum width, and the post focus expansion rate as we increased the emittance in the envelope equations as can be seen in Fig. 3. This suggests that the model used to predict the evolution of the width statistic likewise uses a larger longitudinal emittance than is seen in the N-particle simulation at crossover. As the envelope equations use the longitudinal emittance of the initial warm
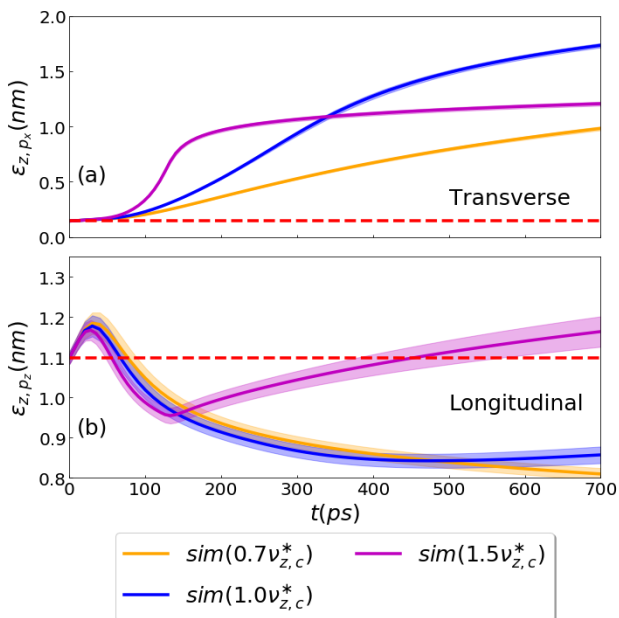
FIG. 6: (rms) Emittance evolution in (a) longitudinal ε_{z,p_z} and (b) transverse ε_{x,p_x} directions for the three different initial chirps, $0.7\ \nu_{z,c}^*$ (orange), $1\ \nu_{z,c}^*$ (blue), and $1.5\ \nu_{z,c}^*$ (magenta). All results were obtained from N-particle simulations. Notice the initial bump in the longitudinal emittance within the first 100 ps; this is driven by disorder induced heating that is not completely resolved by the protocol we used to thermalize the bunches. Further notice that the longitudinal emittance decreases while transverse emittance increases after this point and before the focal point. Also notice that the longitudinal emittance continues to decrease after the focal point for the chirps that are at or below the critical chirp; however, the longitudinal emittance increases at and after the focal point in the crossover regime. This occurs while the transverse emittance continues to increase, so the standard theoretical explanation of heat transfer between the dimensions does not describe this behavior and new theory is required to understand what is going on. This is discussed further in the text.

distribution that is used in the N-particle simulations, this suggests that the longitudinal emittance is in fact decreasing during the N-particle simulations. Tracking the rms emittance of the N-particle simulations, as seen in panel (b) of Fig. 6, confirms that the longitudinal emittance decreases prior to the focal point.

As discussed previously, the conservation of emittance in the envelope equations is a result of the term $\frac{m^2 c^2}{2} \frac{d\varepsilon_{i,p_i}^2}{dt} = s_i^2 s_{p_i,F_i} - s_{i,p_i} s_{i,F_i}$ being 0 in Eq (B2c); conversely, the non-conservation of emittance suggests that this term is non-zero. Currently, there is no theory to predict the value of these terms, but we can use the change in emittance seen in simulations in this term

within the envelope equations to better capture the evolution. Specifically, we replace Eq. (15c) by

$$\frac{d\eta_i^2}{dt} = \frac{d}{dt}\left(\frac{\varepsilon_{i,p_i}^2}{\sigma_i^2}\right) = -\frac{2\gamma_i \eta_i}{m\sigma_i^2} + \frac{1}{\sigma_i^2}\frac{d\varepsilon_{i,p_i}^2}{dt} \qquad (19)$$

in our envelope equations with $\frac{d\varepsilon_{i,p_i}^2}{dt}$ taken from the simulation results. We note that this procedure was originally examined by Sacherer[35]. The spatial width and longitudinal kinetic energy evolution using these envelope equations with the simulation change in emittance squared term can be seen as the dotted lines in Fig. 5. Excellent agreement between these modified envelope equations suggests that varying emittance is the main factor causing the discrepancy between the longitudinal spatial variance and longitudinal kinetic energy evolution of the constant-emittance envelope equations and N-particle simulations. This suggests that if the physics of the covariance terms s_{p_i,F_i} and s_{i,F_i} can be understood and modeled, that we should be able to obtain envelope equations that capture the expected behavior of electron bunches to a high degree of accuracy.

V. DISCUSSION

To provide context for the models examined in this manuscript, we consider an ensemble of N electrons. A full N-particle simulation would find and apply the electric, and if appreciable, the magnetic field at every location. Using the Fast Multipole method[40], this field calculation has a computational order of $\mathcal{O}(N \ln N)$. One computation-saving approximation that can be made is to replace this field calculation by the mean-field although doing such an approximation requires an understanding of the distribution at time t. If we make the further assumption that the distribution is always uniform, this field becomes exceedingly simple and can be parameterized by the standard deviations of the ensemble. Applying this field to the real distribution has a computation order of $\mathcal{O}(N)$ as the field needs to be applied to the N particles at every step. We call this simulation the uniform mean-field real distribution simulation UMFRDS.

As can be seen in Fig. 7, the envelope model exactly reproduces the dynamics of the statistics seen in the UMFRDS despite the computation order of the envelope model having no dependence on N. This is not to say that the UMFRDS and the envelope equations are the same. It is true that the envelope model is able to capture the local velocity spread information that is present in the real distribution through the application of what the accelerator physics field terms normalized rms emittance, and this is the reason for the similarity between the simulation and the model. On the other hand, the UMFRDS has information about the position of the N particles that is lacking in the envelope model. Such
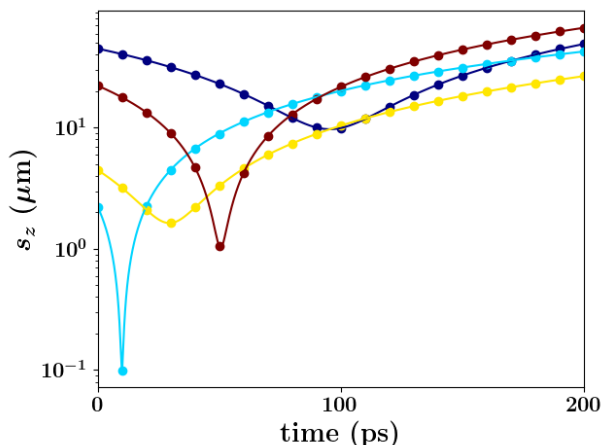
FIG. 7: Plot showing the equivalence between the prediction of the longitudinal standard deviation from the envelope model and the longitudinal standard deviation seen in a N-particle simulation where the force experienced by the particles is assumed to correspond to the uniform mean-field force of an ellipsoid, abbreviated as UMFRDS in the text. The solid lines represent envelope model predictions for 4 different initial parameter choices and the circles correspond to the standard deviation from the UMFRDS.

particle distribution information is important in explaining the shock that occurs in Coulomb explosion[43], the similar shock that arises in a dense Gaussian distribution of electrons within the electron gun[22], or dark matter halos in galaxy formation[44, 45] among other distribution effects. Nonetheless, if such distribution information is not being utilized by the researcher and all that is being examined is some statistic like the width or kinetic energy spread in some dimension, the elimination of any dependence of the algorithm on the number of particles through the use of the envelope model allows us to investigate ensembles of different N with no additional computational cost.

However, this is not the only benefit of the sample perspective we have used here. The statistical kinematics we have used here can be used to show that the evolution of the second order statistics of N particles under a full N-particle simulation can be exactly captured by 9 degrees of freedom assuming there is no force, like a magnetic force, mixing the dimensions. The envelope equation already represents 6 of these degrees of freedom (2 degrees for each of 3 dimensions); the other 3 degrees of freedom can be captured by the evolution of the emittance along the 3 dimensions. Indeed, this is the mathematical foundation underlying Sacherer's observation that *a priori* knowledge of the emittance evolution results in an exact prediction of the statistics evolution[35].

In our work we, are interested in the evolution of these

statistics, and we have presented the envelope equation which assumes conserved emittance. The comparison we have done between the MCE model and the envelope equations can be interpreted as the effect that local velocity spread has on the dynamics of the statistics; however, the envelope model does not currently include stochastic effects. Nonetheless, N-particle simulation have forces that differ from the mean-field assumption implicit in the envelope model and are therefore able to capture such effects. Here, we examined stochastic effects of on the additional 3 degrees of freedom (2 by cylindrical-symmetric assumption) by examining the evolution of the emittance seen in N-particle simulations. Ideally, if we could understand and model the emittance change dynamics, we should eventually be able to exactly (on average) capture the full dynamics of the statistics. Indeed, we find that modeling a distribution that passes through a controlled longitudinal crossover represents an ideal, as well as practically applicable, process where the growth of stochastic effects and hence emittance can be large with a sudden onset of growth at close to the crossover event.

The emittance evolution in Fig. 6, especially within the crossover regime where the emittance in both the longitudinal and transverse directions increase simultaneously, cannot be explained by the standard heat transfer mechanism employed in the literature. We provide some insight into these dynamics here. As can be seen in Fig 6, the longitudinal emittance rapidly increases at the beginning of the simulation followed by a gradual decrease. For the simulations within the crossover regime, there is another rapid increase in the longitudinal emittance close to the focal point. In contrast, the transverse emittance has a rapid increase at the beginning of the simulation followed by a more gradual increase. Notice that there is again a rapid increase in the transverse emittance around the focal point. We emphasize that the rapid increase in the emittance of both directions is almost coincidental — an observation not currently predicted from theory.

Within the literature, there seem to be two macroscopic ideas for the mechanisms involved in this process:

1. *Emittance transfer between degrees of freedom (Heat transfer):* As emittance can be thought as proportional to the square root of the heat times the spatial extent, this heat transfer results in emittance transfer between the degrees of freedom.

2. *Disorder-induced heating (DIH):* DIH in the plasma community describes the heating process during the transition from a disordered state to one which is structured by Coulomb forces.

We point out that these two ideas have previously been described in the literature. Specifically, Reiser's standard book in accelerator physics describes the heat transfer[23], and DIH has phenomenologically been described by Gericke et. al and Maxson et. al[41, 46]. Further, Struckmeier discussed these two ideas in his work on modeling envelope equations with additional Fokker-Planck style random terms[47–49] with slightly different

language and a more mathematical presentation; however, it is not clear to us if these effects are truly stochastic in the manner that should be modeled by Fokker-Planck statistics as Struckmeier has presented. We also point out that these ideas are mechanisms in the language of thermodynamics but do not describe mechanisms in the statistical physics sense as their definition does not lead to any inherent time scale.

Nonetheless, using these two ideas and our own notation, we can phenomenologically explain the emittance evolution seen in Fig. 5. First we define the linear heat along i by $\frac{1}{2m}\eta_i^2$, which can be viewed as the kinetic energy contained in the random fluctuations (from linearity), and we use this measure as a proxy for the heat of the distribution. Recall that the emittance can be written as $\varepsilon_{i,p_i} = \frac{1}{mc}s_i\eta_i$, which can be thought as being proportional to the product of the spatial width and the local momentum width or equivalently the square root of the linear heat. At the beginning of the simulation after the confinement is suddenly removed, potential energy is released through DIH into η_T and η_z while s_T and s_z do not change much. This results in a sudden increase in both ε_{x,p_x} and ε_{z,p_z} that continues even as s_T and s_z begin to significantly change.

Next, the longitudinal emittance begins to decrease as the transverse emittance continues to increase. To understand this phenomenologically, we first consider how the linear heat would change under emittance conserving conditions. As $\eta_z = \frac{\varepsilon_{z,p_z}}{s_z}$, a decrease in s_z would result in an increase in η_z under the assumption of 2D conserved emittance. Likewise, $\eta_x = \frac{\varepsilon_{x,p_x}}{s_x}$, and an increase in s_x would result in a decrease in η_x. In other words if the 2D emittance were conserved, we'd expect the linear heat in the longitudinal direction, i.e. $\frac{1}{2m}\eta_z^2$, to increase, and we'd expect the linear heat in the transverse direction , i.e. $\frac{1}{2m}\eta_x^2$, to decrease. Notice that a strict definition of temperature is not appropriate for our highly non-equilibrium situation, but that this definition of linear heat is still appropriate. As the initial thermalization leads to these two heats being roughly the same, the difference in the heat that develops as the longitudinal dimension contract and the transverse dimension expands would lead to the development of a thermal gradient between the longitudinal and transverse dimensions.

Now in the non-emittance conserving condition, heat can be transferred between the dimensions. So as the bunch focusses, we'd expect a heat transfer from the hotter longitudinal to the cooler transverse direction — that is, the evolution of simulated η_z would be expected to be smaller than the η_z we'd obtain from the emittance preserving envelope equations; conversely, we'd expect the simulated η_x to be larger than the theory η_x. This is precisely what is seen in Fig. 8. This in turn results in the longitudinal emittance decreasing while the transverse emittance increases, which is precisely what happens in the bounce-back regime.

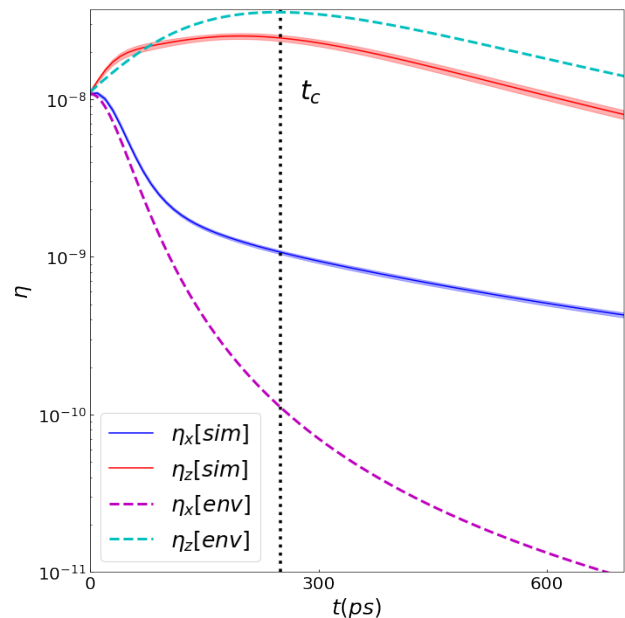However, once the bunch is in the crossover regime, there is actually an increase in the longitudinal emit-

FIG. 8: Comparison of the evolution of the parameter η between the envelope equations (dashed lines) and the N-particle simulations (solid lines). The line and shaded area around the N-particle simulation lines represent the mean (solid line) $\pm$ the standard deviation of 90 simulations. The vertical dotted line indicates the focal point. Notice that the emittance conserving model over-predicts η_z and under-predicts η_x. This is partially due to the heat being transferred between the dimensional modes – a mechanism that is not captured by the emittance conserving envelope equations.

tance in Fig 6. This phenomenon could be explained from this perspective by postulating that a second period of DIH occurs near the focal point when the bunch is within the crossover regime. That is, if the particles crossover, they are forced into a highly non-equilibrium state that rapidly relaxes releasing heat into the bunch thus increasing the emittance.

VI. CONCLUSIONS

In this work, we have examined the longitudinal crossover of electron bunches with uniform ellipsoidal profiles focused by a linear chirp as is typical of the propagation of a probing electron bunch in an ultrafast electron diffraction/microscope system. We employed several analytic techniques to model the space charge dynamics of the bunch, the first of which is an extension of Grech et. al's mean-field theory which utilizes ordinary differential equations for the ellipsoid transverse and longitudinal sizes to describe the bunch evolution. Analysis of this mean-field model leads to the identification of a longitudinal critical chirp. This critical chirp separates two

regimes for particle trajectories in this model: bounce-back, where the particles reverse their direction at the bunch waist, and crossover, where the bunch experiences a singularity with a width of zero. This singularity is only seen in the zero emittance case and is analogous to the unphysical caustics seen in gravitation analysis[44, 45]; while such effects are unphysical as zero emittance for any ensemble is unphysical, this classification is nonetheless useful. We showed that time can be scaled by the initial plasma frequency, and by defining a dimensionless critical chirp the zero-emittance model behavior depends only on the initial aspect ratio. The evolution of bunches with the same initial geometry then differ only by the time scale determined by the bunch's plasma frequency.

We examined the problem through the statistical formulation of envelope equations by building on Sacherer's statistical analysis of the cylindrical KV-envelope equations that are well known in the accelerator physics community. We argue that our model is widely applicable to the evolution of statistics of any ensemble far beyond the application we examined here. In other work, we recently presented a statistical perspective on the envelope equations that we are calling the sample perspective[38]. We showed that the statistical envelope equations for three dimensional systems are identical to, up to a constant we determine in Appendix B 2, the Analytical Gaussian formalism that is well known in the ultrafast electron microscopy community; the envelope equations from the sample perspective are more general and have a more straightforward derivation than Michalik and Sipe's integral approach. Again, we note that this derivation is not limited to accelerator applications but can be applied to any statistical analysis.

It should also be noted that the majority of analyses of envelope equations in the accelerator physics community is in cylindrical coordinates due to the predominance of accelerators with continuous beams or which have bunches with very large prolate aspect ratios. In contrast in the UEM field, the bunch near the source and at the longitudinal focal point is a highly oblate ellipsoid, or pancake, and a fully 3D description is required. The advantage of our analysis is that the envelope model captures such 3D effects, and one of the clear results is that the zero emittance envelope equation clearly reveals the critical chirp phenomenon separating the bounce back and crossover regimes at the bunch focal point described above.

However, our envelope model assume linear forces that conserve the statistical measure of emittance. Fortunately, investigating the evolution of emittance in N-particle simulations should In specific, we provided a qualitative description of the emittance growth and transfer observed in Fig. 6 by elucidating three mechanisms: (i) Disorder induced heating, or relaxation from a highly non-equilibrium state, which converts potential energy into kinetic energy and naturally leads to a rapid growth in emittance. This effect occurs at the start of our simulations and near the focal point. (ii) Transfer

of "heat", which is equal to kinetic energy fluctuations, from the hotter direction to the colder direction. (iii) The fact that the kinetic energy fluctuations increase along a direction that is compressed and decrease along a direction that is expanding. This last mechanism allows for one direction to become hotter than the other colder, e.g. when the longitudinal direction is being compressed it becomes hotter while the transverse direction cools down as it expands. The system then tries to equilibrate by transferring heat from the hot direction to the cold direction. We note that the longitudinal emittance, the longitudinal momentum fluctuations and the longitudinal bunch size are related by $\epsilon_{z,p_z} = s_z \eta_z$ and that the longitudinal non-equilibrium heat is given by $\eta_z^2/2m$. Therefore dynamics of the heat, or momentum fluctuations, and the dynamics of the emittance may be different due to the s_z factor as is evident when comparing Figs. 6 and 8. Further model development and characterization of the randomness inherent in such consistent effects is an exciting avenue for future insight into particle-particle, finite-size effects.

Appendix

Appendix A: A statistical perspective on accelerator physics

The following section is primarily aimed toward accelerator physicists and may not be of interest to students and experts outside of that discipline. In the accelerator field, the Courant-Snyder parameters, which are related to the covariance parameters we use in the main text[50], are often used to describe bunch motion. In this section, we do 2 things: 1.) discuss why our problem cannot use the Courant-Snyder parameters and 2.) show the statistical kinematics equations in terms of Courant-Snyder parameters.

1. Trace space and longitudinal normalized rms emittance

In order to explain why we think Courant-Snyder parameters are not appropriate for this problem and to introduce our notation, we first discuss the paraxial assumption. The paraxial assumption is necessary for the use of trace space, and this assumption actually has two imbedded assumptions that we describe now. Consider N particles indexed by j. The velocity of the jth particle,

$$\vec{v}_j = v_{x,j}\hat{x} + v_{y,j}\hat{y} + v_{z,j}\hat{z}, \qquad (A1)$$

is assumed to be primarily in the z direction, i.e. $v_{z,j} >> v_{x,j}, v_{y,j}$. As a result, $\beta_j = \frac{\sqrt{\vec{v}_j \cdot \vec{v}_j}}{c} \approx \frac{v_{z,j}}{c}$. A second assumption is then made — all particles are assumed to have the same β, i.e. $\beta_j = \beta$ for all j. As a result, the velocity of the jth particle can be approximated as

$$\vec{v}_j \approx c\beta \left(\frac{v_{x,j}}{c\beta}\hat{x} + \frac{v_{y,j}}{c\beta}\hat{y} + \hat{z} \right). \qquad (A2)$$

Moreover, the coordinate system that accelerator physicists use when doing analyses has some subtle innovations; namely, the coordinate system can be thought of as moving with the center of mass of the ensemble of particles so that the z-direction is always along the direction with the largest velocity even when the trajectory of the particles is bent. In more strict mathematical language, the expected path of the bunch is parameterized by a curvilinear parameter, s by convention, and the coordinate system is a tangent space associated with the center of mass's location.

As v_z is always along s, we have $\frac{ds}{dt} = v_z \approx c\beta$. Notice $v_{x,j} = \frac{dx_j}{dt} = \frac{dx_j}{ds}\frac{ds}{dt} = c\beta x_j'$ where $'$ indicates the derivative with respect to s. Likewise, $v_{y,j} = c\beta y_j'$. Subbing this into Eq. (A3) we get

$$\vec{v}_j \approx c\beta \left(x_j'\hat{x} + y_j'\hat{y} + \hat{z} \right) \qquad (A3)$$

and accelerator physicists call the coordinates (x, x') and (y, y') 2D trace spaces analogous (and related to) the standard (x, p_x) and (y, p_y) 2D phase spaces.

In the trace treatment, β_j and hence the Lorentz factor γ_j are assumed to be approximately the same for all particles j — that is, accelerator physicists treat these parameters as statistical scalars. Defining $\gamma = \gamma_j$, this allows us to pull both β and γ out of the statistics. As $p_{x,j} = mc\gamma_j\beta_j x_j' \approx mc\gamma\beta x_j'$, it follows $s_{x,p_x} \approx mc\gamma\beta s_{x,x'}$ and $s_{p_x,p_x} \approx m^2c^2\gamma^2\beta^2 s_{x,x'}$, at least to first order, which is typically a very good approximation in the standard situation examined by accelerator physicists. As a consequence, $\varepsilon_{x,p_x}^2 = \frac{1}{m^2c^2}(s_x^2 s_{p_x}^2 - s_{x,p_x}^2) = \gamma^2\beta^2\epsilon_{x,x'}^2$ where

$$\epsilon_{x,x'}^2 = s_x^2 s_{x'}^2 - s_{x,x'}^2 \qquad (A4)$$

represents the statistical area occupied by the distribution in (x,x') trace space and is known as the rms emittance. Analogous expressions can be found in y.

However, the z-statistics need to be treated with care as the standard approximation leads to an unphysical 0 normalized rms emittance in the longitudinal rms emittance, which we denote by ε_z. Treating $p_{z,j} = p_z$ results in $s_{p_z,p_z} = 0$ and $s_{z,p_z} = 0$. Therefore, the longitudinal normalized rms emittance, $\varepsilon_{z,p_z} = \frac{1}{mc}\sqrt{s_z^2 s_{p_z}^2 - s_{z,p_z}^2}$, defined in Reiser's central textbook "The Theory and Design of Charged Particle Beams"[23], is likewise 0 to first order. This is unphysical, so first order is not sufficient to describe the longitudinal normalized rms emittance.

Instead, accelerator physicists treat the longitudinal emittance to second order. Treating $p_{z,j} = p_z + \delta_j$ where δ_j is the small deviation from p_z for the jth particle with $\bar{\delta} = 0$ and uncorrelated with momentum, i.e. $s_{p_z,\delta} = 0$, we get

$$s_{z,z} = s_z^2 \qquad (A5a)$$
$$s_{z,p_z} = s_{z,\delta} \qquad (A5b)$$
$$s_{p_z,p_z} = s_\delta^2 \qquad (A5c)$$

Further assuming $s_{z,\delta} = 0$, the longitudinal normalized rms emittance becomes $\varepsilon_{z,p_z} = \frac{1}{mc}s_z s_\delta$. Normalizing the longitudinal normalized rms emittance by the same factor that the transverse normalized rms emitttances are normalized by to give the transverse rms emittances, i.e. $\gamma\beta$, we get the longitudinal rms emittance

$$\varepsilon_z = \frac{1}{\gamma m\beta c}s_z s_\delta. \qquad (A6)$$

The above equation is not directly used by accelerator physicists; instead they measure the longitudinal emittance in terms of the energy spread and the duration of the pulse. Note, though, that such an expression is obtainable from Eq. (A6). Notice that the duration of the pulse, Δt, is proportional to $\Delta t \propto \frac{s_z}{\beta c}$. Further, notice that the total kinetic energy of the pulse is given by

$$E^2 = c^2 \sum_{j=1}^{N} (\vec{p}_j \cdot \vec{p}_j + m^2c^2)$$
$$\approx c^2 \sum_{j=1}^{N} (p_{z,j}^2 + m^2c^2)$$
$$= Ns_{p_z}^2 c^2 + N\bar{p}_z^2 c^2 + Nm^2c^4$$
$$\approx Ns_\delta^2 + E_{CoM}^2. \qquad (A7)$$

where $E_{CoM}^2 \approx N\bar{p}_z^2 c^2 + Nm^2c^4$ is the energy associated with the center of mass motion that can be shown to be approximately true unless $\frac{p_z}{mc} \approx 1$. Further approximating $(\Delta E)^2 = s_E^2 \approx \frac{1}{N}(E^2 - E_{com})$ we obtain $s_\delta = \Delta E$. As a result we have

$$\varepsilon_z \propto \Delta t \Delta E \qquad (A8)$$

under these approximations giving a statistical motivation for the standard accelerator physics approximation.

The problem with analyzing longitudinal focussing within the standard framework is that $s_{z,p_z} = s_{z,\delta}$ needs to be non-zero in order for the bunch to focus — a central premise of the main text. There are a couple of problems with generalizing the approximations made by the accelerator community when such a covariance is present. First, $s_{z,\delta}$ becomes important in the determination of the kinetic energy therefore invalidating the relation $\varepsilon_z \propto \Delta t \Delta E$ as ΔE becomes an overestimate of s_δ. Even if this problem can be handled, the treatment of the second order effect in the z-direction needs to be properly propagated throughout the derivation. This includes the definition of the rms emittance as the approximation for the z-velocity of the jth particle becomes $v_{z,j} \approx \bar{v}_z + \frac{1}{\gamma m} \frac{s_{z,p_z}}{s_z}^2 z$ and the definition of the trace variable will thus depend on z and trace space cannot be used.

We do not see a solution to this problem and therefore have considered non-relativistic phase-space where the mathematics is much simpler and generally applicable. As the Courant-Snyder parameters are defined in trace space, they cannot be used to describe our problem.

2. The statistical kinematics reformulated in Courant-Snyder parameters

If the longitudinal dynamics can be ignored, we can reformulate the statistical kinematic equations in terms of trace space parameters. This is known in the accelerator field as Wangler's theorem is essentially derived from these considerations[51]; however, we have not seen these considerations derived in terms of Courant-Snyder parameters. We do the Courant-Snyder derivation in this section.

We first point out that the relation between the second order statistics of the trace parameters and the Courant-Snyder parameters is well known[50]:

$$\hat{\beta} = \frac{s_x^2}{\epsilon_{x,x'}} \tag{A9a}$$

$$\hat{\alpha} = -\frac{s_{x,x'}}{\epsilon_{x,x'}} \tag{A9b}$$

$$\hat{\gamma} = \frac{s_{x',x'}}{\epsilon_{x,x'}}. \tag{A9c}$$

From this and Eq. (A4), we have the relation

$$1 = \hat{\beta}\hat{\gamma} - \hat{\alpha}^2. \tag{A9d}$$

The rms emittance parameter is often used instead of one of the 3 Courant-Snyder parameters. The contstraint in Eq. (A9d) might make it appear as if there are only two degrees of freedom, but we point out that the Courant-Snyder parameters are dependent on the rms emittance. Therefore, Eqs. (A9a) - (A9c) actually have four degrees of freedom, and the constraint brings the degrees of freedom back to 3. Despite only having three degrees

of freedom, we will write the following equations in all 4 parameters, $\hat{\beta}$, $\hat{\alpha}$, $\hat{\gamma}$, and $\epsilon_{x,x'}$, as is standard to keep the notation from becoming unwieldy.

Moreover, the choice of 3 parameters is no accident. The number of parameters for the kth order statistical kinematics can be counted using combinatorics, i.e the degrees of freedom are given by $\binom{k+1}{1} = k + 1$, which is 3 for second order statistical kinematics. If these three parameters are given, the evolution of these parameters are exactly known if the details of the force are exactly known as implied by Sacherer[35]. Namely taking the derivative of Eqs.(A9a) - (A9d) we obtain

$$\frac{d\hat{\beta}}{ds} = -2\hat{\alpha} - \frac{\hat{\beta}}{\epsilon_{x,x'}} \frac{d\epsilon_{x,x'}}{ds} \tag{A10a}$$

$$\frac{d\hat{\alpha}}{ds} = -\hat{\gamma} - \frac{s_{x,x''}}{\epsilon_{x,x'}} - \frac{\hat{\alpha}}{\epsilon_{x,x'}} \frac{d\epsilon_{x,x'}}{ds} \tag{A10b}$$

$$\frac{d\hat{\gamma}}{ds} = 2\frac{s_{x',x''}}{\epsilon_{x,x'}} - \frac{\hat{\gamma}}{\epsilon_{x,x'}} \frac{d\epsilon_{x,x'}}{ds} \tag{A10c}$$

$$0 = \frac{d\hat{\beta}}{ds}\hat{\gamma} + \hat{\beta}\frac{d\hat{\gamma}}{ds} - 2\hat{\alpha}\frac{d\hat{\alpha}}{ds}$$
$$= 2\hat{\beta}\frac{s_{x',x''}}{\epsilon_{x,x'}} + 2\hat{\alpha}\frac{s_{x,x''}}{\epsilon_{x,x'}} - \frac{2}{\epsilon_{x,x'}}\frac{d\epsilon_{x,x'}}{ds} \tag{A10d}$$

where

$$x'' = \frac{d}{ds}x'$$
$$\approx \frac{d}{ds}\left(\frac{1}{\gamma\beta mc}p_x\right)$$
$$= -\frac{\gamma'}{\gamma}x' - \frac{\beta'}{\beta}x' + \frac{1}{\gamma\beta mc}\frac{1}{\frac{ds}{dt}}\frac{d}{dt}p_x$$
$$= -\frac{\gamma^2\beta'}{\beta}x' + \frac{1}{\gamma\beta^2 mc^2}F_x \tag{A11}$$

and F_x is the x-component of the force measured in the lab frame. Solving Eq. (A10d) for the rms emittance, we get

$$\frac{d\epsilon_{x,x'}}{ds} = \hat{\beta}s_{x',x''} + \hat{\alpha}s_{x,x''} \tag{A12}$$

We note that Eq. (A12) has been in the accelerator literature for some time[51]. We emphasize that only three of the derivatives are needed to describe the system, but $\frac{d\hat{\gamma}}{ds}$ is given above to show the derivation of the derivative of the rms emittance. Subbing Eq. (A11) into the formula for covariance we get

$$s_{x,x''} = -\frac{\gamma^2\beta'}{\beta}s_{x,x'} + \frac{1}{\gamma\beta^2 mc^2}s_{x,F_x}$$
$$= \frac{\gamma^2\beta'\hat{\alpha}\epsilon_{x,x'}}{\beta} + \frac{1}{\gamma\beta^2 mc^2}s_{x,F_x} \tag{A13a}$$

$$s_{x',x''} = -\frac{\gamma^2\beta'}{\beta}s_{x'}^2 + \frac{1}{\gamma\beta^2 mc^2}s_{x',F_x}$$
$$= -\frac{\gamma^2\beta'\hat{\gamma}\epsilon_{x,x'}}{\beta} + \frac{1}{\gamma\beta^2 mc^2}s_{x',F_x} \tag{A13b}$$

Plugging these expression into Eqs. (A10a), (A10b, and (A12), we get

$$\frac{d\hat{\beta}}{ds} = -2\hat{\alpha} - \frac{\hat{\beta}}{\epsilon_{x,x'}}\frac{d\epsilon_{x,x'}}{ds} \tag{A14a}$$

$$\frac{d\hat{\alpha}}{ds} = \hat{\gamma} + \frac{\gamma^2\beta'}{\beta}\hat{\alpha} - \frac{1}{\gamma\beta^2 mc^2}\frac{s_{x,F_x}}{\epsilon_{x,x'}} - \frac{\hat{\alpha}}{\epsilon_{x,x'}}\frac{d\epsilon_{x,x'}}{ds} \tag{A14b}$$

$$\begin{aligned}
\frac{d\epsilon_{x,x'}}{ds} &= \hat{\beta}\left(-\frac{\gamma^2\beta'\hat{\gamma}\epsilon_{x,x'}}{\beta} + \frac{1}{\gamma\beta^2 mc^2}s_{x',F_x}\right) \\
&\quad + \hat{\alpha}\left(\frac{\gamma^2\beta'\hat{\alpha}\epsilon_{x,x'}}{\beta} + \frac{1}{\gamma\beta^2 mc^2}s_{x,F_x}\right) \\
&= \left(-\gamma^2\frac{\beta'}{\beta}\epsilon_{x,x'} + \frac{1}{\gamma\beta^2 mc^2}\left(\hat{\alpha}s_{x,F_x} + \hat{\beta}s_{x',F_x}\right)\right).
\end{aligned} \tag{A14c}$$

In the cases when the functional form of the x-force can be exactly provided for all time, the statistical kinematics predict the evolution of the Courant-Snyder parameters as long as the paraxial assumption is valid. Alternatively, models employ assumed forces can be developed under certain conditions similar to the fluid and envelope models presented in Reiser's book[23]. Such prediction can even be done for normalized rms and rms emittance, which we are preparing for publication.

Notice that if $F_x = 0$ but the beam is accelerated so $\beta' > 0$, then Eq. (A14c) gives $\frac{d\epsilon_{x,x'}}{ds} < 0$. This phenomenon is well known and is called adiabatic damping[23]. In general, though, forces can be incorporated into these equations via the term dependent on $\hat{\alpha}s_{x,F_x} + \hat{\beta}s_{x',F_x}$ which we will call A below for the sake of brevity. Typically, F_x is parameterized as some function of x. For instance, if the x-force is linear in x, i.e. $F_x = kx$ where k is not a function of x or x' (but can be a function of t), then this term becomes

$$\begin{aligned}
A &= \hat{\alpha}s_{x,kx} + \hat{\beta}s_{x',kx} \\
&= k\hat{\alpha}s_{x,x} + k\hat{\beta}s_{x',x} \\
&= k(\hat{\alpha}\hat{\beta} - \hat{\beta}\hat{\alpha})\epsilon_{x,x'} \\
&= 0 \tag{A15}
\end{aligned}$$

which is a another well known result[23]. Alternatively, when the x-force has non-linear components, the A term contributes to the rms emittance change. These two facts provided the motivation for Wangler to express the emittance growth of a non-uniform bunch in relation to the difference between the energy of the initial non-uniform and the final uniform-like state[51].

Appendix B: Relation between AG model and envelope equations

1. Equivalence of notation

We denote the parameters used in the AG model by superscript AG. Further, we identify that the AG model's parameters are related to the covariance parameters we use in this manuscript by

$$\sigma_i^{AG} = s_i^2 \tag{B1a}$$

$$\gamma_i^{AG} = s_{i,p_i} \tag{B1b}$$

$$\eta_i^{AG} = s_{p_i}^2 - \frac{s_{i,p_i}^2}{s_i^2} \tag{B1c}$$

Subbing these parameters into our ODE we obtain

$$\frac{d\sigma_i^{AG}}{dt} = \frac{2}{m}\gamma_i^{AG} \tag{B2a}$$

$$\frac{d\gamma_i^{AG}}{dt} = \frac{1}{m}\left(\frac{\gamma_i^{AG2}}{\sigma_i^{AG}} + \eta_i^{AG}\right) + s_{i,F_i} \tag{B2b}$$

$$\frac{d\eta_i^{AG}}{dt} = -\frac{2\gamma_i^{AG}\eta_i^{AG}}{m\sigma_i^{AG}} + \frac{2}{\sigma_i^{AG}}\left(\sigma_i^{AG}s_{p_i,F_i} - \gamma_i^{AG}s_{i,F_i}\right) \tag{B2c}$$

This system of equations differs from Michalik and Sipe's published system of ODE's in two ways: 1.) in Eq. (B2b) we have s_{i,F_i} instead of Michalik and Sipe's $\frac{1}{4\pi\epsilon_0}\frac{Ne^2}{6\sqrt{\pi}\sigma_i}L_i\left(\frac{s_z}{s_T}\right)$ where $L_i\left(\frac{s_z}{s_T}\right)$ is an integral we examine in detail in the next section and 2.) we include $\frac{2}{\sigma_i^{AG}}\left(\sigma_i^{AG}s_{p_i,F_i} - \gamma_i^{AG}\sigma_{i,F_i}\right) = \frac{1}{s_i^2}\frac{dm^2c^2\epsilon_{i,p_i}^2}{dt}$ where $\epsilon_{i,p_i}^2 = \frac{1}{m^2c^2}(s_i^2p_i^2 - s_{i,p_i}^2) = \sigma_i^{AG}\eta_i^{AG}$ in Eq. (B2c) which Michalik and Sipe treat as 0. As noted by our notation, the latter additional term represents the effect of the change of emittance on the bunch evolution. We believe that that Michalik and Sipe's assumption of self-similar evolution leads to this term vanishing and hence emittance being conserved. This assumption is not strictly true as we have recently shown that the Gaussian distribution does not evolve self-similarly under the Coulomb force[22]; nonetheless, self-similarity may be a reasonable assumption in order to approximate spatial statistics.

However, we can obtain the published AG model under the assumption that the force can be written as $F_i = ki$ where $k = \frac{1}{4\pi\epsilon_0}\frac{Ne^2}{6\sqrt{\pi}s_i^{AG3/2}}L_i\left(\frac{s_z}{s_T}\right)$. This results in $s_{p_i,F_i} = k_i\gamma_i^{AG}$ and $s_{i,F_i} = k_is_i^{AG}$ reducing the emittance change term to zero, i.e. $\frac{1}{s_i^2}\frac{dm^2c^2\epsilon_{i,p_i}^2}{dt} = \frac{2}{\sigma_i^{AG}}\left(\sigma_i^{AG}k_i\gamma_i^{AG} - \gamma_i^{AG}k_i\sigma_i^{AG}\right) = 0$. Moreover, Eq.

(B2b) becomes

$$\frac{d\gamma_i^{AG}}{dt} = \frac{1}{m}\left(\frac{\gamma_i^{AG^2}}{\sigma_i^{AG}} + \eta_i^{AG}\right) + k_i\sigma_i^{AG}$$

$$= \frac{1}{m}\left(\frac{\gamma_i^{AG^2}}{\sigma_i^{AG}} + \eta_i^{AG}\right) + \frac{1}{4\pi\epsilon_0}\frac{Ne^2}{6\sqrt{\pi}s_i^{AG^{1/2}}}L_i\left(\frac{s_z}{s_T}\right).$$
(B3)

This is the published form of the AG model[31]

2. AG model and the uniform assumption

In the main text we argue that conservation of emittance leads to an assumed linear force on the ensemble particles. Here, we show that the evolution of the AG model is equivalent to the envelope equations assuming a uniform distribution up to a constant. Specifically, we show that the force portion of Eq. (B2b) obtained by Michalik and Sipe by integration techniques is the same up to a constant to the analogous term obtained by using the mean-field force within a uniform ellipsoid. Knowing that $s_{i,k_ii} = k_is_i^2$ we infer slopes of the force for Michalik and Sipe to

$$k_i^{MS} = \frac{1}{4\pi\epsilon_0}\frac{Ne^2}{6\sqrt{\pi}s_i^3}L_i\left(\frac{s_z}{s_T}\right)$$
(B4)

where

$$L_z(a) = \frac{3a^2}{a^2-1}\left(aL(a)-1\right)$$
(B5a)

$$L_T(a) = \frac{3}{2}\left(L(a) + \frac{a^2L(a)-a}{1-a^2}\right)$$
(B5b)

$$L(a) = \begin{cases} \frac{arcsin\sqrt{1-a^2}}{\sqrt{1-a^2}}, & 0 \le a \le 1 \\ \frac{\ln(a+\sqrt{1-a^2})}{\sqrt{a^2-1}}, & 1 \le a \end{cases}$$
(B5c)

On the other hand, the slopes of the force for a uniform ellipsoid with $s_x = s_y$ can be written as

$$k_x^{unif} = \frac{1}{4\pi\epsilon_0}\frac{3Ne^2}{10\sqrt{5}s_x^3}\beta\left(1,\frac{s_z}{s_x}\right)$$
(B6a)

$$k_z^{unif} = \frac{1}{4\pi\epsilon_0}\frac{3Ne^2}{10\sqrt{5}s_x^3}\beta\left(\frac{s_x}{s_z},\frac{s_x}{s_z}\right)$$
(B6b)

(B6c)

where

$$\beta(a,b) = \int_0^\infty \frac{1}{(1+u)^{3/2}\sqrt{a^2+u}\sqrt{b^2+u}}du$$
(B7)

A derivation of this is presented in our recent work[38]. The comparison between these two models comes down to how Eq. (B5) and Eq. (B7) compare. Specifically, letting $a = \frac{s_z}{s_x}$ as it is in Eq. (B5), we need to evaluate $\beta(1,a)$ and $\beta(\frac{1}{a},\frac{1}{a})$ and then compare the two slopes.

We start the evaluation of the transverse relevant integral:

$$\beta(1,a) = \int_0^\infty \frac{1}{(1+u)^2\sqrt{a^2+u}}du$$

$$= \frac{cos^{-1}(a) - a\sqrt{1-a^2}}{(1-a^2)^{3/2}}$$

$$= \begin{cases} \frac{sin^{-1}(\sqrt{1-a^2})-a\sqrt{1-a^2}}{(1-a^2)^{3/2}}, & 0 \le a \le 1 \\ \frac{i\ln(a+\sqrt{a^2-1})-a\sqrt{1-a^2}}{(1-a^2)^{3/2}}, & a \ge 1 \end{cases}$$

$$= \begin{cases} \frac{sin^{-1}(\sqrt{1-a^2})}{\sqrt{1-a^2}(1-a^2)} - \frac{a}{1-a^2}, & 0 \le a \le 1 \\ \frac{\ln(a+\sqrt{a^2-1})}{\sqrt{a^2-1}(1-a^2)} - \frac{a}{1-a^2}, & a \ge 1 \end{cases}$$

$$= \frac{L(a)-a}{1-a^2}$$

$$= L(a) + \frac{a^2L(a)-a}{1-a^2}$$

$$= \frac{2}{3}L_T(a)$$
(B8)

Thus, our uniform integrals in the transverse direction differs by a factor of $\frac{2}{3}$ from the AG model's Gaussian integrals in the transverse direction.

Next we evaluate the longitudinally relevant integral:

$$\beta\left(\frac{1}{a},\frac{1}{a}\right) = \int_0^\infty \frac{1}{(1+u)^{3/2}(\frac{1}{a^2}+u)}$$

$$= \frac{2}{\frac{1}{a^2}-1} - \frac{2sec^{-1}\left(\frac{1}{a}\right)}{(\frac{1}{a^2}-1)^{3/2}}$$

$$= \frac{2a^2}{(1-a^2)}\left(1 - a\frac{cos^{-1}(a)}{\sqrt{1-a^2}}\right)$$

$$= \begin{cases} \frac{2a^2}{(a^2-1)}\left(a\frac{sin^{-1}(\sqrt{1-a^2})}{\sqrt{1-a^2}} - 1\right), & 0 \le a \le 1 \\ \frac{2a^2}{(a^2-1)}\left(a\frac{i\ln(a+\sqrt{a^2-1})}{\sqrt{1-a^2}} - 1\right), & a \ge 1 \end{cases}$$

$$= \begin{cases} \frac{2a^2}{(a^2-1)}\left(a\frac{sin^{-1}(\sqrt{1-a^2})}{\sqrt{1-a^2}} - 1\right), & 0 \le a \le 1 \\ \frac{2a^2}{(a^2-1)}\left(a\frac{\ln(a+\sqrt{a^2-1})}{\sqrt{a^2-1}} - 1\right), & a \ge 1 \end{cases}$$

$$= \frac{2}{3}L_z(a)$$
(B9)

Again we see that the same factor is present between the integrals in the longitudinal direction, which is reassuring.

Putting this factor of $\frac{2}{3}$ into the comparison of the slopes, we see that the slopes are related by

$$\frac{k_i^{MS}}{k_i^{unif}} = \frac{5\sqrt{5}}{6\sqrt{\pi}} \approx 1.05$$
(B10)

So we see that these two models differ in their forces only by roughly 5%. This small difference in the linear force should result in either model adequately capturing the dynamics of either uniform or Gaussian evolution if the changes in emittance, that should be more prevalent

in the Gaussian model, are ignored. Furthermore, even this difference may be absorbed by the assumed number of particles when fitting parameters. That is, in both modes N needs to be set. As the only term that depends on N is the force, setting the N in the uniform envelope equations 5% larger than the N in the AG model will result in the same exact solution. So in essence the AG model is identical to the uniform envelope equation but with a slightly adjusted number of particles.

---

[1] E. Hall, S. Stemmer, H. Zheng, Y. Zhu, and G. Maracas, *Future of Electron Scattering and Diffraction*, Tech. Rep. (US Department of Energy, Washington, DC (United States), 2014).

[2] E. Lessner, X. Wang, and P. Musumeci, *Report of the Basic Energy Sciences Workshop on the Future of Electron Sources*, (SLAC National Accelerator Laboratory, 2016).

[3] R. D. Miller, Annual review of physical chemistry **65**, 583 (2014).

[4] R. D. Miller, Science **343**, 1108 (2014).

[5] R. Srinivasan, V. Lobastov, C.-Y. Ruan, and A. Zewail, Helvetica Chimica Acta **86**, 1761 (2003).

[6] J. R. Dwyer, C. T. Hebeisen, R. Ernstorfer, M. Harb, V. B. Deyirmenjian, R. E. Jordan, and R. Dwayne Miller, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **364**, 741 (2006).

[7] A. H. Zewail, Annu. Rev. Phys. Chem. **57**, 65 (2006).

[8] C.-Y. Ruan, Y. Murooka, R. K. Raman, R. A. Murdick, R. J. Worhatch, and A. Pell, Microscopy and Microanalysis **15**, 323 (2009).

[9] G. Sciaini and R. D. Miller, Reports on Progress in Physics **74**, 096101 (2011).

[10] A. A. Ischenko and S. A. Aseyev, in *Advances in Imaging and Electron Physics*, Vol. 184 (Elsevier, 2014) pp. 1–26.

[11] J. Xu, C. I. Blaga, P. Agostini, and L. F. DiMauro, Journal of Physics B: Atomic, Molecular and Optical Physics **49**, 112001 (2016).

[12] W. E. King, G. H. Campbell, A. Frank, B. Reed, J. F. Schmerge, B. J. Siwick, B. C. Stuart, and P. M. Weber, Journal of Applied Physics **97**, 111101 (2005).

[13] A. H. Zewail, science **328**, 187 (2010).

[14] D. A. Plemmons, P. K. Suri, and D. J. Flannigan, Chemistry of Materials **27**, 3178 (2015).

[15] A. Feist, N. Bach, N. R. da Silva, T. Danz, M. Möller, K. E. Priebe, T. Domröse, J. G. Gatzmann, S. Rost, J. Schauss, *et al.*, Ultramicroscopy **176**, 63 (2017).

[16] Z. Tao, H. Zhang, P. Duxbury, M. Berz, and C.-Y. Ruan, Bulletin of the American Physical Society **56** (2011).

[17] Z. Tao, H. Zhang, P. Duxbury, M. Berz, and C.-Y. Ruan, Journal of Applied Physics **111**, 044316 (2012).

[18] J. Portman, H. Zhang, Z. Tao, K. Makino, M. Berz, P. Duxbury, and C.-Y. Ruan, Applied Physics Letters **103**, 253115 (2013).

[19] J. Portman, H. Zhang, K. Makino, C. Ruan, M. Berz, and P. Duxbury, Journal of Applied Physics **116**, 174302 (2014).

[20] J. Portman, H. Zhang, K. Makino, C. Ruan, M. Berz, and P. Duxbury, Advances in Imaging and Electron Physics **191**, 117 (2015).

[21] J. Williams, F. Zhou, T. Sun, Z. Tao, K. Chang, K. Makino, M. Berz, P. Duxbury, and C.-Y. Ruan, Structural Dynamics **4**, 044035 (2017).

[22] B. Zerbe, X. Xiang, C.-Y. Ruan, S. Lund, and P. Duxbury, Physical Review Accelerators and Beams **21**, 064201 (2018).

[23] M. Reiser, *Theory and Design of Charged Particle Beams* (John Wiley & Sons, New York, 1994).

[24] O. J. Luiten, S. B. vanderGeer, M. J. deLoos, F. B. Kiewiet, and M. J. vanderWiel, Physical review letters **93**, 094802 (2004).

[25] B. J. Siwick, J. R. Dwyer, R. E. Jordan, and R. J. Dwayne Miller, Journal of Applied Physics **92**, 1643 (2002).

[26] B.-L. Qian and H. E. Elsayed-Ali, Journal of Applied Physics **91**, 462 (2002).

[27] B. W. Reed, Journal of Applied Physics **100**, 034916 (2006).

[28] S. Collin, M. Merano, M. Gatri, S. Sonderegger, P. Renucci, J.-D. Ganiere, and B. Deveaud, Journal of applied physics **98**, 094910 (2005).

[29] C. C. Lin, L. Mestel, and F. H. Shu, The Astrophysical Journal **142**, 1431 (1965).

[30] M. Grech, R. Nuter, A. Mikaberidze, P. Di Cintio, L. Gremillet, E. Lefebvre, U. Saalmann, J. M. Rost, and S. Skupin, Physical Review E **84**, 056404 (2011).

[31] A. Michalik and J. Sipe, Journal of applied physics **99**, 054908 (2006).

[32] A. Michalik and J. Sipe, Journal of Applied Physics **105**, 084913 (2009).

[33] J. A. Berger and W. A. Schroeder, Journal of Applied Physics **108**, 124905 (2010).

[34] I. Kapchinskij and V. Vladimirskij, CERN, Scientific Information Service, Geneva , 274 (1959).

[35] F. J. Sacherer, IEEE Transactions on Nuclear Science **18**, 1105 (1971).

[36] O. Anderson, Part. Accel. **21**, 197 (1987).

[37] P. M. Lapostolle, T. Wangler, and A. Lombardi, *A model of non-linear space-charge forces in a charged particle beam*, Tech. Rep. (CM-P00059509, 1993).

[38] B. S. Zerbe and P. M. Duxbury, Submitted (2020).

[39] L. D. Landau, *The classical theory of fields*, Vol. 2 (Elsevier, 2013).

[40] Z. Gimbutas and L. Greengard, Communications in Computational Physics **18**, 516 (2015).

[41] D. Gericke and M. Murillo, Contributions to Plasma Physics **43**, 298 (2003).

[42] S. Plimpton, Journal of Computational Physics **117**, 1 (1995).

[43] A. E. Kaplan, B. Y. Dubetsky, and P. L. Shkolnikov, Physical review letters **91**, 143401 (2003).

[44] P. Zukin and E. Bertschinger, Physical Review D **82**, 104045 (2010).

[45] P. Zukin and E. Bertschinger, Physical Review D **82**, 104044 (2010).

[46] J. Maxson, I. Bazarov, W. Wan, H. Padmore, and C. Coleman-Smith, New Journal of Physics **15**, 103024

(2013).

[47] J. Struckmeier, Part. Accel. **45**, 229 (1994).

[48] J. Struckmeier, Physical Review E **54**, 830 (1996).

[49] J. Struckmeier, Physical Review Special Topics-

Accelerators and Beams **3**, 034202 (2000).

[50] J. Buon, *Beam phase space and emittance*, Tech. Rep. (Paris-11 Univ., 1992).

[51] T. Wangler, K. Crandall, R. Mills, and M. Reiser, IEEE Transactions on Nuclear Science **32**, 2196 (1985).