

# Space-charge distortion of transverse profiles measured by electron-based ionization profile monitors and correction methods

D. Vilsmeier\*

Johann Wolfgang Goethe-University Frankfurt, 60323 Frankfurt am Main, Germany

M. Sapinski and R. Singh

GSI Helmholtz Centre for Heavy Ion Research, 64291 Darmstadt, Planckstraße 1, Germany



(Received 20 November 2018; published 13 May 2019)

Measurements of transverse profiles using ionization profile monitors (IPMs) for high brightness beams are affected by the electromagnetic field of the beam. This interaction may cause a distortion of the measured profile shape despite strong external magnetic field applied to impose limits on the transverse movement of electrons. The mechanisms leading to this distortion are discussed in detail. The distortion itself is described by means of analytic calculations for simplified beam distributions and a full simulation model for realistic distributions. Simple relation for minimum magnetic field scaling with beam parameters for avoiding profile distortions is presented. Further, application of machine learning algorithms to the problem of reconstructing the actual beam profile from distorted measured profile is presented. The obtained results show good agreement for tests on simulation data. The performance of these algorithms indicate that they could be very useful for operations of IPMs on high brightness beams or IPMs with weak magnetic field.

DOI: 10.1103/PhysRevAccelBeams.22.052801

## I. INTRODUCTION

The principle of beam profile measurement using ionization profile monitors (IPMs) is seemingly very simple. The beam particles ionize the residual gas. The products of the ionization—electrons or ions—are extracted towards a position-sensitive detector using the guiding electric field (also referred to as “clearing” or “external” field) provided by electrodes. The distribution of the particles on the detector ideally corresponds to a projection of the transverse distribution of the beam density. This simple and straightforward principle, illustrated in Fig. 1, is nondestructive for the beam and thus an appealing application for synchrotrons and storage rings. In this context IPMs are installed in many hadron accelerators [1] and even investigated for usage on electron machines [2].

However already for medium-intensity beams the movement of electrons or ions is easily affected by the transient beam electric field. Because of their lower mass, electrons are removed much faster from the influence of bunch fields, however the effect of these fields on their trajectories is

proportionally larger. Usually ion-based devices are preferred due to simpler assembly (no need of ion-trapping [4]), smaller impact of fringe fields of neighbouring magnets and lack of background signal from electrons generated from various other sources [5]. However the influence of beam fields plays a major role for high-brightness beams and thus electron-based IPMs are preferred over ion-based devices, mainly because the electron movement can be easily confined to small gyroradius by applying external magnetic field. Another argument in

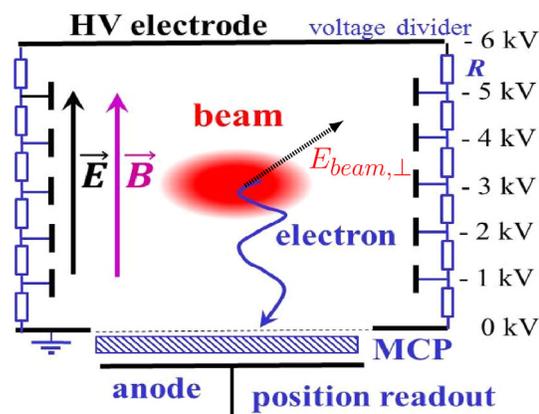


FIG. 1. Sketch of an IPM together with read-out system. The electric and magnetic guiding fields are aligned, perpendicular to the beam direction as well as to the read-out system [3].

\*d.vilsmeier@gsi.de

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

favor of electron-based devices is the fact that they are more suitable for time-resolved measurements because the electrons have much smaller time-of-flights towards the detector.

Distortions of the measured beam profile using ion-based IPMs are either mitigated by increasing the electric guiding field or corrected by empirical mathematical models [6–9]. For operation of electron-based IPMs with magnetic field it was first proposed to tune the magnetic field strength such that the electrons perform exactly one gyration between the beam and the detector [10]. However this method does not work for high-brightness beams, in which electrons get a significant momentum transfer not only transverse to the detector plane but also in the direction towards the detector. Therefore increasing the magnetic field in order to reduce the magnitude of gyroradius is the typical counteraction against the distortion of the measured beam profiles. This is also because profile correction methods, similar to the ones developed for ion-based IPMs, do not exist for electron-based devices. One reason for that is the more complex particle movement in the presence of a combination of electric and magnetic fields. This movement and therefore the profile distortion, strongly depend on the beam fields and hence on bunch charge, transverse, and longitudinal bunch shape as well as bunch spacing.

The first beam space-charge induced effects in an electron-based IPM operated with magnetic guiding field were observed in the Large Hadron Collider (LHC) at beam energies of 4 TeV and magnetic guiding field of 0.2 T [11]. Following studies have analyzed this phenomenon and estimated a magnetic field strength of 1 T required to suppress profile distortions [12–14]. However such large field strengths pose a technological as well as financial challenge, especially because of the relatively large gap between the poles needed to fit the detector vacuum chamber.

Besides beam space-charge induced profile distortion there are various other phenomena which can have an effect on measured profiles. These are briefly discussed below however not addressed any further throughout this study: (1) Ionization by bunch field. At extreme field strengths the gas ionization cross section is modified due to the Stark shift of energy levels and at even higher fields the gas gets ionized by the collective bunch electric field [15], rather than by interaction with single beam particles. (2) Burnout of the rest gas. Under some conditions, the ionization rate of the rest gas can be higher than the replenishment of gas in the volume of the beam leading to nonproportionality between beam density and amount of ionization events. This can be counterbalanced by, for instance, using gas jet [16]. (3) Guiding field nonuniformity. If the electric and magnetic guiding fields are nonuniform, it will additionally impact the resulting particle trajectories [17]. Also if the guiding fields are

not perfectly aligned with respect to each other this introduces an additional disturbance. (4) Gas ionization from synchrotron radiation and the influence of wake fields on particle trajectories may constitute additional sources of profile distortion. However these effects have not been observed or studied up to now.

Some of the mentioned effects emerge from the influence of the particle beam and depend mainly on the amplitude of the beam electric field and the duration of the corresponding pulse. An overview of various effects in dependence on these two parameters is shown in Fig. 2 alongside the beam parameters of several hadron accelerators, the most powerful laser system up-to-date (LFEX), 3rd generation light source (ALBA) and advanced free electrons laser (SwissXFEL). The peak fields in waist of LFEX laser beam reaches  $1 \times 10^7$  MV m<sup>-1</sup>, well beyond the limit of gas ionization by bunch electric field. In comparison, the highest field in an accelerator—a focused SwissXFEL beam—will be reaching about  $2 \times 10^4$  MV m<sup>-1</sup>. All hadron machines are currently below 10 MV m<sup>-1</sup> but the bunch lengths are about 3 orders of magnitude larger than those of LFEX and XFEL. Figure 2 also indicates two effects related to beam space-charge interaction: electrons being trapped by the beam fields exceeding the electric guiding field and electron velocities reaching into the relativistic regime.

Depending on the acquisition system in use, additional distortion of the measured distributions may occur. The electron or ion detection is usually based on amplification of the signal using multichannel plate (MCP) and either conversion of resulting electrons to light using a phosphor screen followed by a camera or direct conversion to electric signals using anode-strips. Such systems may suffer from

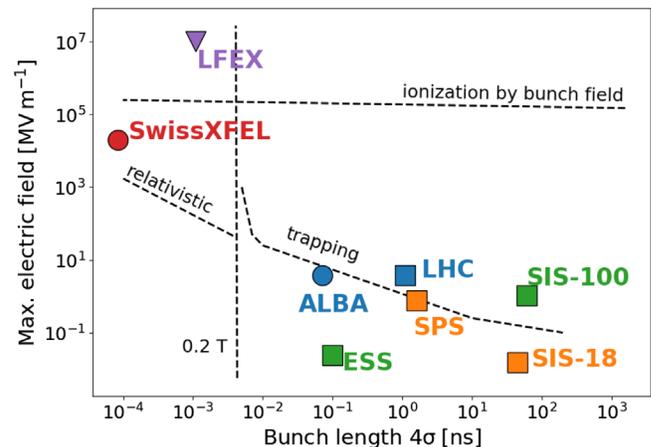


FIG. 2. Amplitude and duration of transient beam electric fields generated in various scientific apparatus. Approximate boundaries indicating appearance of phenomena like electron trapping or relativistic regime are shown. The estimates were done for a magnetic guiding field of 0.2 T. The dashed vertical line indicates the gyrioperiod for electrons moving in the 0.2 T magnetic field.

nonuniformity of MCP gain, Phosphor deterioration or point-spread function of the optical system positioned after the Phosphor. Recent developments focus on use of hybrid silicon pixel detector instead of MCP/Phosphor or MCP/anode-strip combination and have shown to sidestep some of these issues [18].

The present study focuses on application of electron-based IPMs in high-brightness hadron machines such as the Large Hadron Collider (LHC), the Super Proton Synchrotron (SPS) at CERN or the SIS-100 under construction at FAIR. For these machines the beam-space induced profile distortion becomes a prominent factor that potentially hinders the successful operation of electron-based IPMs. This type of profile distortion as well as possible mitigation strategies are the focus of the present study.

## II. PROFILE DISTORTION

In the following sections we will focus on profile distortion effects related to the electron movement which is governed by their interaction with the present electromagnetic fields. For the scope of this contribution, we assume perfectly uniform and aligned guiding fields and do not consider gas target and acquisition system related effects which were mentioned in the previous section. The coordinate system used throughout this study is  $x$ , along the measured profile (horizontal),  $y$ , towards the acquisition system (vertical) and  $z$ , along the beam (longitudinal).

For illustration purposes we refer in this section to the data obtained for an example case which has been studied by means of simulations. The corresponding parameters are given in Table I and correspond to a possible LHC flat top configuration (nominal value of bunch charge is about  $1.3 \times 10^{11}$ , but charges above  $2 \times 10^{11}$  are expected for HL-LHC). The LHC beam case is well studied [12–14, 19–21]. As another example we will shortly discuss the SIS-100 IPM system for proton beam at flat top.

TABLE I. Beam and device parameters for the example case which corresponds to a possible LHC flat top configuration.

Parameter	Value
Beam particle type	Protons
Energy/u	6.5 TeV
Bunch population	$2.1 \times 10^{11}$ ppb
Bunch length ( $4\sigma_z$ )	0.9 ns
Bunch width ( $1\sigma_x$ )	270 $\mu\text{m}$
Bunch height ( $1\sigma_y$ )	360 $\mu\text{m}$
Electrode distance	85 mm
Applied voltage	4 kV
Magnetic field	0.2 T
Number of sim. electrons	100 000
Time step size	0.3125 ps

### A. Effects related to electron movement

The electron movement is initialized by the ionization process and further influenced by the interaction with the guiding fields and the electromagnetic field of the particle beam (*space-charge interaction*). The relevant effects can be summarized by the following three aspects:

(i) *Ionization momenta*—The initial momenta of electrons, obtained during the ionization process, clearly influence their further trajectories. Due to the interaction with the magnetic guiding field an initial momentum component along the beam axis results in a transverse displacement of the subsequent gyromotion while a transverse momentum component results in a longitudinal displacement. This effect becomes more pronounced for small guiding field strengths as well as for small beam sizes where the magnitude of the gyromotion increases relative to the beam dimensions. For relativistic beams the electrons are mainly ejected transverse to the beam direction [22].

(ii) *Space-charge interaction*—The interaction of electrons with the electric field of the particle beam may significantly alter their trajectories, resulting in noticeable distortions of measured profiles. This interaction cannot be treated analytically without introducing significant approximations and hence must be studied by the means of numerical simulations. As this effect correlates with the magnitude of the beam electric field, it becomes especially significant for high intensity and high energy beams. However also small magnetic guiding fields may not be sufficient to counteract the beam field interaction and hence provoke an increase in gyroradii.

(iii) *Gyromotion*—The electrons perform a spiral movement under the influence of the external magnetic guiding field. Depending on the magnitude of the gyroradius an electron may experience a significant displacement with respect to its center of gyration. This effect is especially significant for small beam sizes since the magnitude of the displacement increases relative to the beam size.

Often the magnetic guiding field is sufficient for suppressing profile distortion. However in extreme cases the space-charge interaction can become so strong, with electric fields of up to a few  $\text{MV m}^{-1}$  (typical extraction fields are around  $50 \text{ kV m}^{-1}$ ), that electrons are (a) trapped inside the space-charge region since the beam electric field outweighs the electric guiding field around the center of the bunch, and (b) forced on significantly different trajectories, resulting in a vast increase of gyroradii. Since the above effects depend on the various parameters of the particle beam it is often not obvious to determine an appropriate strength for the magnetic guiding field which allows to sufficiently suppress the gyroradius increase.

Having in mind the above-mentioned effects we may subdivide the IPM volume into two regions:

(i) *Space-charge region*—This is the region in the vicinity of the beam where the interaction with the beam's electromagnetic field has a major influence. The gyrovelocity of

electrons are subject to a perpetual oscillation induced by the  $\vec{E} \times \vec{B}$  drift along the beam and the magnitude of this oscillation is altered when the beam center approaches and recedes from the positions of the electrons. Because of the complex shape of the electromagnetic fields the electron motion exhibits other electromagnetic drifts as well, for example *polarization drift* due to the time dependence of the beam electric field [23]. For estimating the dimensions of this region one can consider the  $\vec{E} \times \vec{B}$  drift velocity as an indicator for space-charge action. The region in which this  $\vec{E} \times \vec{B}$  drift velocity is more than 1% of the undisturbed gyrovelocity, is of the order of a few millimeters.

(ii) *Detector region*—This is the region close to the acquisition system where the beam fields diminish to a negligible magnitude. This means the electron movement are solely subject to the electric and magnetic guiding fields and hence the electrons perform a pure gyromotion planar to the detector while being accelerated toward it. The characteristics of this gyromotion (gyroradius and gyrocenter position) depend on the previous beam space-charge interaction.

Note that as the bunch recedes from the electrons' positions the space-charge region will shrink in time. While this region can be quite large during the bunch center passing, the transverse electric field quickly diminishes according to the bunch's line density. Therefore also electrons that are ionized near the bunch's center will eventually end up performing a pure gyromotion as in the detector region. Figure 3 shows the bunch electric field as it decreases towards the detector ( $y$ ) and along the beam axis ( $z$ ) and hence illustrates the separation in space-charge and detector region. While the two dependencies are shown

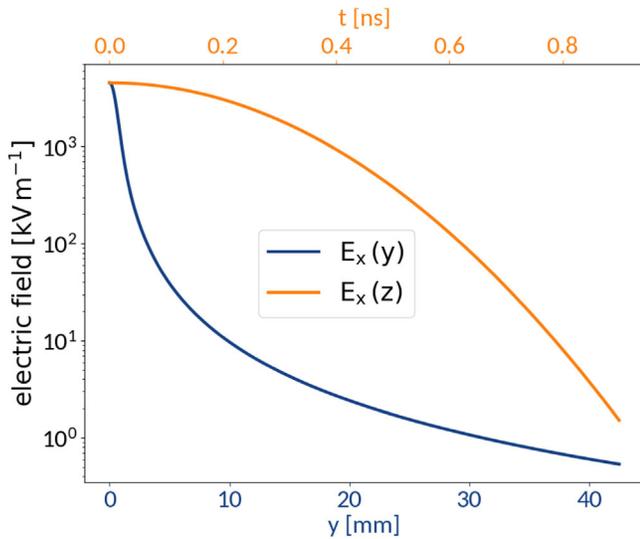


FIG. 3. Evolution of the transverse bunch electric field  $E_x$  at  $x = \sigma_x$  in direction toward the detector ( $y$ ) and during beam passage ( $z$ ,  $t$ ), for the parameters in Table I. The  $x$ -position is fixed to  $1\sigma_x$  and  $y$ - and  $z$ -position are fixed to zero if not varied.

separately, their collective effect results in a much stronger decrease as the bunch recedes from the electrons' positions and hence the effective space-charge region is smaller than indicated by the single dependencies. This is because the beam velocity is significantly larger than the electrons' velocities along the beam axis and hence the spatial dependence ( $z$ ) can be thought of as a time-dependence ( $t$ ).

Considering an electron's *initial* position during ionization and *final* position during detection,  $x_0$  and  $x_f$  respectively, the overall displacement  $\Delta x = x_f - x_0$  can be ascribed to three different effects: (1)  $\Delta x_1$ —Displacement of the electron's gyrocenter  $x_0^c$  with respect to the initial position, due to its initial velocity along the beam:  $\Delta x_1 = x_0^c - x_0$ . Electrons that are ejected transverse to the longitudinal beam axis, in horizontal direction, merely suffer from a displacement of their gyrocenter along that axis, i.e., not affecting their horizontal position which determines the transverse profile. If they are ejected along the beam however there will be a corresponding displacement transverse to it. This type of displacement is typically around a few tens of microns and can be assessed with differential ionization cross sections which usually depend on the projectiles charge state and energy. In relativistic cases, transverse ejection dominates and so the resulting displacement is mainly along the beam [22]. (2)  $\Delta x_2$ —Displacement of the space-charge altered gyrocenter  $x_s^c$  with respect to the initial gyrocenter:  $\Delta x_2 = x_s^c - x_0^c$ . This shift is found to be induced by the magnetic field of the beam. In addition the longitudinal electric field of the beam induces an  $\vec{E} \times \vec{B}$  drift in horizontal direction which adds up as an additional displacement of the gyrocenter. However for relativistic beams the exposure to the longitudinal field is short and hence this drift tends to be negligible. The net effect of this type of displacement is around a few microns even for high beam currents. (3)  $\Delta x_3$ —Displacement of the final position with respect to the stable gyrocenter due to the nature of the gyromotion:  $\Delta x_3 = x_f - x_s^c$ . The electrons perform a gyromotion in the plane above the detector and in case this motion spans over multiple elements of the detecting system, the electrons might be detected on any of them. If the moment of detection is not correlated to the phase of the gyromotion then the resulting displacement is random. Since this type of displacement is proportional to the resulting gyroradii it depends on the magnitude of the preceding space-charge interaction and can be around some hundreds of microns up to a few millimeter.

Figure 4 shows an example trajectory which exhibits the different electromagnetic drifts, the separation in space-charge and detector region, as well as the different kind of displacements  $\Delta x_{\{1,2,3\}}$  mentioned above. The increase in gyroradius is also clearly visible. The maximum polarization drift distance, corresponding to the parameters in Table I, is estimated to be  $852 \mu\text{m}$  as further discussed in Appendix A.

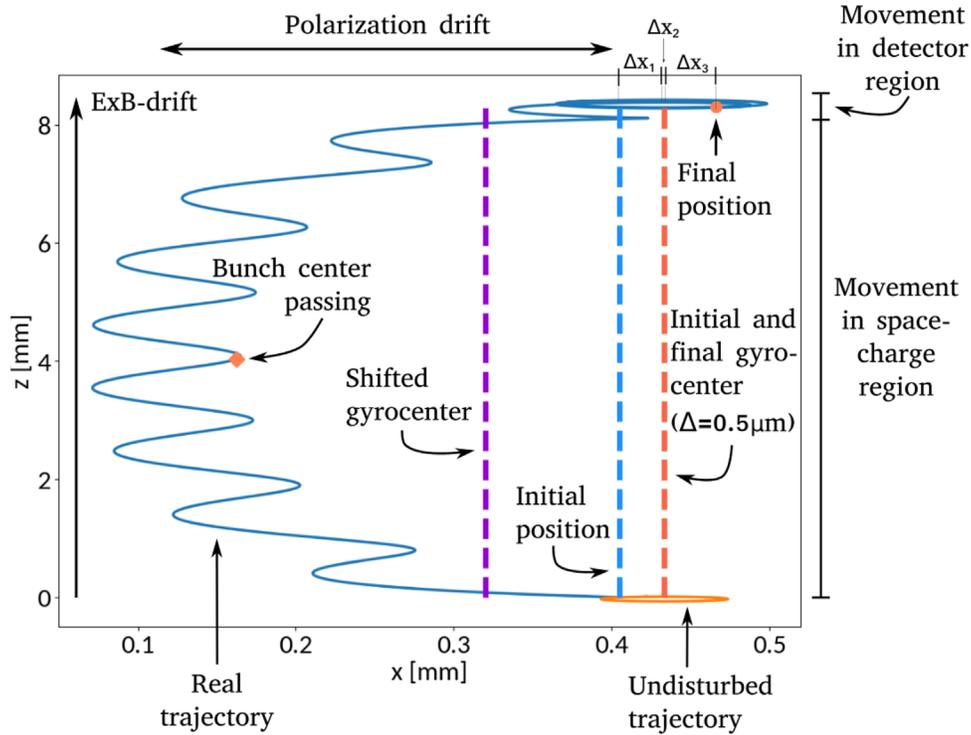


FIG. 4. Example trajectory obtained for  $\mathbf{r}_0 = (1.5\sigma_x, 0, 0)$ ,  $\mathbf{v}_0 = 10^6 \cdot (-1, 0, 1) \text{ m s}^{-1}$  and generated  $2\sigma_z$  before the bunch center (beam parameters corresponding Table I). The term “undisturbed trajectory” refers to the case if there were no space-charge interaction, i.e., only pure gyromotion. Because the initial and final gyroradius is about  $100 \mu\text{m}$  the corresponding part of the trajectory appears as a flat ellipse due to large scale of the  $z$ -drift. The “shifted gyrocenter” is induced by the transverse electric field of the beam and acts as a starting point for the polarization drift; this shift is described by equation Eq. (6). “Bunch center passing” indicates the moment when the electron and the bunch center are aligned with respect to  $z$ -position. The “initial” and “final gyrocenter” are separated by a distance of  $0.5 \mu\text{m}$ . The labels “Movement in detector/space-charge region” refer to the period that is spent in each of the regions, not the spatial domain, as the vertical axis indicates movement along the beam axis. There is no transverse  $\vec{E} \times \vec{B}$  drift because the simulation was configured with  $E_z = 0$ .

These effects only consider displacements in horizontal direction, along the measured profile. In fact the electrons experience an additional significant displacement in longitudinal direction, along the beam, due to the  $\vec{E} \times \vec{B}$  drift induced by the transverse electric field of the beam  $E_x$ . However such drifts can be neglected in case the physical situation is similar along the beam axis (i.e., uniform guiding fields, similar gas pressure). In typical setups the size of the readout system is much smaller than the corresponding area of the surrounding IPM field box and it is centered therein. Hence these conditions are satisfied.

Figure 5 shows the velocity increase of electrons which is induced by the space-charge interaction and consequentially leads to a deformation of measured profiles as shown by Fig. 6. The mean of initial velocities is  $9.94 \times 10^5 \text{ m s}^{-1}$  which corresponds to a gyroradius of  $28.3 \mu\text{m}$ . The total displacement of an electron is compound of the three above-mentioned stages:  $\Delta x = \Delta x_1 + \Delta x_2 + \Delta x_3$ . Regarding the transverse  $\vec{E} \times \vec{B}$  drift due to nonzero longitudinal field component  $E_z$  one can consider the fact that the transverse field is scaled with relativistic factor of  $\gamma \approx 7000$ . Since the

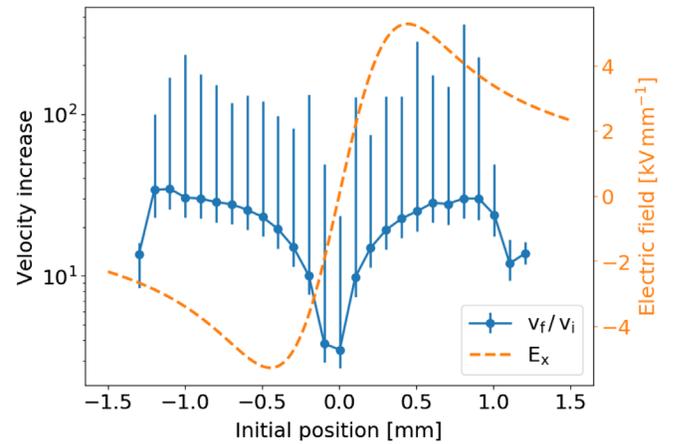


FIG. 5. Velocity increase, given as the ratio of velocity at detection and velocity at ionization, plotted against the initial horizontal position of electrons. Dots denote average values while vertical bars indicate 1 standard deviation. The transverse electric field in  $x$ -direction, measured at  $y = z = 0$  at the center of the bunch, is overlaid. Beam parameters corresponding to Table I.

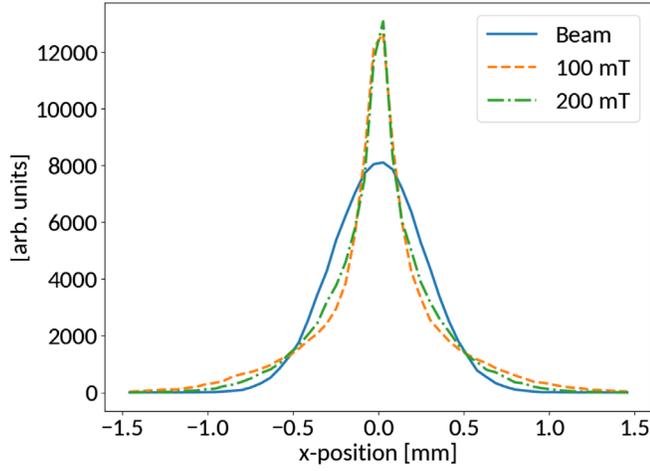


FIG. 6. Simulated example profiles for beam parameters corresponding to Table I using magnetic guiding field strengths of 100 mT and 200 mT, respectively. The beam standard deviation is  $270 \mu\text{m}$  and the resulting standard deviations of the simulated profiles are  $366 \mu\text{m}$  and  $305 \mu\text{m}$ , respectively.

$\vec{E} \times \vec{B}$  drift velocity is proportional to the electric field strength and the longitudinal drifts typically are of the order of a few millimeter the expected  $\vec{E} \times \vec{B}$  drift distance transverse to the beam axis is expected to be at least  $\gamma^{-1}$  times smaller, resulting in transverse  $\vec{E} \times \vec{B}$  drift distances in the submicrometer regime.

### B. Description for uniform beam distributions

Electron trajectories are mainly influenced by the transverse electric field of the beam and the magnetic guiding field. This interaction alters the trajectories in a complex way which cannot be predicted analytically for any realistic beam distribution without further assumptions.

In order to get an analytic estimate of the interaction we assume a simplified case in which the longitudinal electric field is zero and the transverse electric field is approximated by a linearly increasing field  $E_x(x) = E \cdot x$  where  $E < 0$  (this corresponds to the electric field inside a uniformly charged cylinder bunch of positive ions). For a Gaussian bunch shape this approximation is also well applicable at the center of the bunch as can be seen from Fig. 5 (neglecting the  $y$ - and  $z$ -dependence of the field). We also consider no bunch magnetic field since it is small compared to the magnetic guiding field  $B$  which is acting in  $y$ -direction. The electric guiding field  $E_y$  is acting in  $y$ -direction as well and hence has no influence on the motion in the  $xz$ -plane. In this simplified scenario the particle motion in the  $xz$ -plane is obtained as:

$$x(t) = \frac{v_{x0}}{\Omega} \sin(\Omega t) + \frac{\bar{q}}{\Omega^2} (Bv_{z0} - Ex_0) \cos(\Omega t) + \left(1 + \frac{\bar{q}E}{\Omega^2 B}\right) x_0 - \frac{\bar{q}B}{\Omega^2} v_{z0}, \quad (1)$$

$$z(t) = -\frac{\bar{q}B}{\Omega^2} v_{x0} \cos(\Omega t) + \frac{\bar{q}^2 B}{\Omega^3} (Bv_{z0} - Ex_0) \sin(\Omega t) + \left[ \left( \frac{E}{B} + \frac{\bar{q}E^2}{\Omega^2 B} \right) x_0 - \frac{\bar{q}E}{\Omega^2} v_{z0} \right] t + z_0 + \frac{\bar{q}B}{\Omega^2} v_{x0}, \quad (2)$$

$$\dot{x}(t) = v_{x0} \cos(\Omega t) - \frac{\bar{q}}{\Omega} (Bv_{z0} - Ex_0) \sin(\Omega t), \quad (3)$$

$$\dot{z}(t) = \frac{\bar{q}B}{\Omega} v_{x0} \sin(\Omega t) + \frac{\bar{q}^2 B}{\Omega^2} (Bv_{z0} - Ex_0) \cos(\Omega t) + \left( \frac{E}{B} + \frac{\bar{q}E^2}{\Omega^2 B} \right) x_0 - \frac{\bar{q}E}{\Omega^2} v_{z0}, \quad (4)$$

where we have used the following abbreviations ( $q$  is the elementary charge and  $m$  is the electron mass):

$$\bar{q} \equiv \frac{q}{m} \quad (5a)$$

$$\Omega^2 \equiv \bar{q}^2 B^2 - \bar{q}E > 0, \quad \Omega = \sqrt{\Omega^2}. \quad (5b)$$

Equation (1) shows that the particle performs an oscillating movement in  $x$ -direction while Eq. (2) shows a similar oscillation as well as an  $\vec{E} \times \vec{B}$  drift in  $z$ -direction. The absolute value of the gyrovelocity is given by  $|\mathbf{v}_{xz}(t)| = \sqrt{\dot{x}(t)^2 + \dot{z}(t)^2}$ . From Eq. (3) and Eq. (4) we can see that this velocity expression contains terms  $\sin^2(\Omega t)$ ,  $\cos^2(\Omega t)$ , and  $\sin(\Omega t) \cdot \cos(\Omega t)$  which are  $\pi$ -periodic as well as terms  $\sin(\Omega t)$  and  $\cos(\Omega t)$  which are  $2\pi$ -periodic. Figure 7 shows the gyrovelocity evolution for three example particles subject to the realistic electric field of a Gaussian bunch. The expected twofold periodicity is clearly visible while an additional damping effect occurs due to the longitudinal field dependence when the beam recedes from the particles' positions. The minima and maxima of the velocities become apparent when considering the two-dimensional trajectories  $(x(t), z(t))$ . At one turn-around point the velocities from  $\vec{E} \times \vec{B}$  drift and the oscillating movement are aligned and add up, resulting in a large velocity while at the other turn-around point they are opposite resulting in a smaller velocity (or even a backward drift). Such behavior can as well be observed from simulations as described below. Figure 7 shows the velocities of various particles for different starting points  $x_0$  subject to the realistic field of a Gaussian charge distribution. One can observe a variation of the amplitude modulation as well as different periods of the oscillation depending on the starting point. This is due to the nonlinear shape of the Gaussian-bunch electric field for positions farther away from the bunch center (compare the electric field plot in Fig. 5).

Analyzing Eq. (1) to Eq. (4) we can infer various parameters of the resulting electron motion.

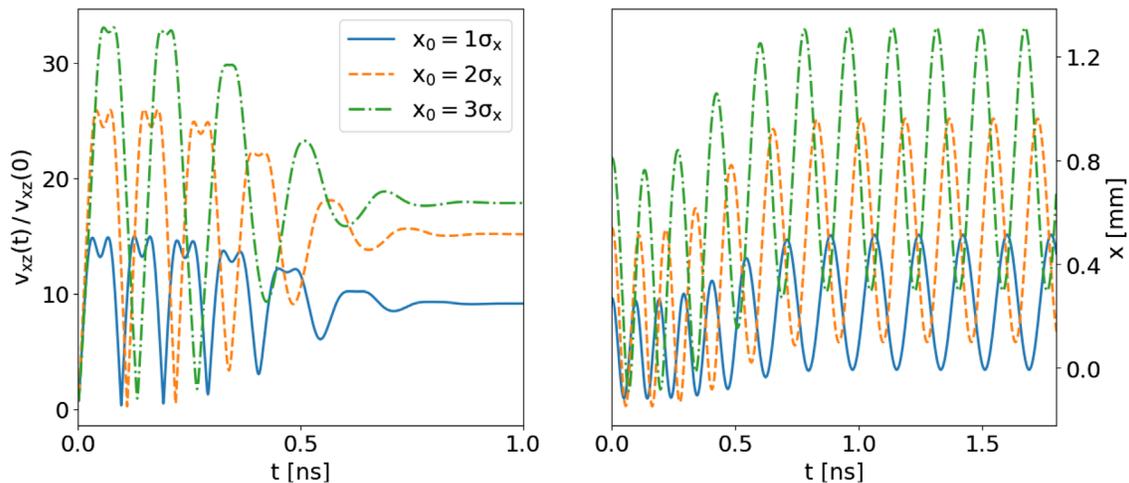


FIG. 7. Left: Gyrovelocity evolution for different starting positions  $\mathbf{r}_0 = (x_0, 0, 0)$ . Initial velocity is  $\mathbf{v}_0 = 10^6 \sqrt{2}^{-1} \cdot (1, 1, -1)$  m s<sup>-1</sup> and electrons are generated  $\sigma_z/2$  before the bunch center (beam parameters corresponding to Table I). Right: The corresponding x-position evolution which turns into pure gyromotion for  $t \gtrsim 1$  ns.

### 1. Shift of gyration center

From Eq. (1) one can read the center of oscillation

$$x_c = \left(1 + \frac{\bar{q}E}{\Omega^2}\right)x_0 - \frac{\bar{q}B}{\Omega^2}v_{z0} \quad (6)$$

which contains an additional term corresponding to the transverse bunch electric field. This shift however is compensated due to the time-dependence of the electric field and the corresponding polarization drift as shown in Appendix B. The same effect can be observed from simulations as well. This results in the initial and final gyrocenter to be aligned with respect to each other as shown in Fig. 4. Note that a displacement of the gyrocenter with respect to the initial position  $x_0$  however occurs due to a nonzero initial velocity component along the beam  $v_{z0}$ . This shift is not compensated for by any other effects.

### 2. Displacement due to gyromotion

In the detector region, without the influence of the beam electric field, the electrons perform a pure gyromotion with gyroradius proportional to the planar velocity. Because of the velocity oscillations in the space-charge region, described by Eq. (3) and Eq. (4) as well as Fig. 7, the electrons are likely to end up with an increased velocity in the detector region and consequentially with an increased gyroradius. Figure 8 shows the distribution of gyrovelocity change. Around 90% of electrons end up with an increased gyrovelocity due to space-charge interaction and hence end up with increased gyroradii.

The resulting gyromotion implies possible displacements to positions different from the center of the gyromotion. This effect is discussed in detail in [13,14].

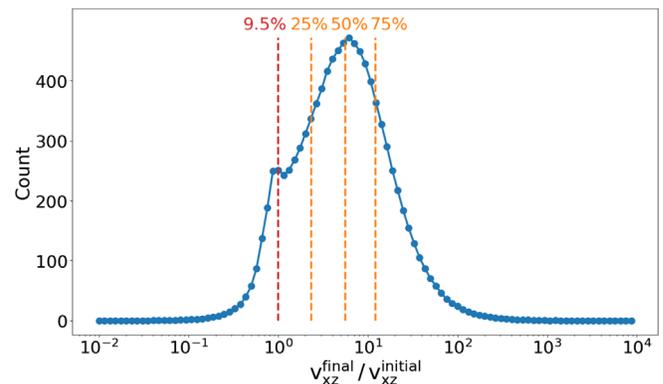


FIG. 8. Distribution of gyrovelocity ratios, final over initial velocity, due to space-charge interaction (parameters corresponding to Table I).

### 3. Time-of-flight

In order to obtain an estimate for the final gyroradius of electrons we need to derive an estimate for their time-of-flight until they reach away from the influence of disturbing beam fields. Since the above analytic considerations are valid only inside the bunch, we set the space-charge region as the region inside the bunch and the detector region as the region outside the bunch. This assumption neglects modifications of the particles' velocities outside the bunch while in reality they may remain affected by the bunch electric field at these positions as the space-charge region may extend farther out. A more complex description would be to treat the decreasing part of the field outside the bunch by another linear field model, similarly to the one inside, and to reuse the above derived equations for that case by applying appropriate coordinate transformations. However since this approach is significantly more complex we stick

with the former approach as it turns out to be sufficient as indicated by the following simulation results.

For estimating the time-of-flights one cannot simply rely on the time the bunch needs to recede from the electron but one needs to consider the electron leaving the bunch volume (i.e., the space-charge region) in the transverse direction as well. In order to obtain a more accurate estimate for this time-of-flight in the vertical direction we need to take the bunch electric field and the electric guiding field into account. We do so by simplifying the longitudinal time dependence by an exponential dependency instead of the Gaussian dependency, in order to relax the explicit time dependence. The spatial dependence is linear as before. The corresponding equation of motion is

$$\ddot{y}(t) = \bar{q}E_g - \bar{q}Ey \exp\left(-\frac{|t|}{\sigma_z}\right) \quad (7)$$

where  $E_g$  is the electric guiding field strength.

For a detailed derivation of the resulting motion  $(y(t), \dot{y}(t))$  consider Appendix C.

In order to estimate the time-of-flights we need to find the roots of the functions  $f_1(t) = \sqrt{x(t)^2 + y(t)^2} - R$

where  $R$  is the radius of the cylinder bunch and  $x(t)$  is given by Eq. (1) and  $f_2(t) = z(t) - (t - 4\sigma_z)\beta c$  where  $\beta = v/c$ , with  $c$  the speed of light; whatever value of  $t$  is smaller determines the moment when the electron leaves the bunch. The final parameters are then obtained by plugging this value back into the analytic velocity equations Eq. (3) and Eq. (4). For more details on the computation consider Appendix D.

Figure 9 shows the gyroradius distributions computed the analytic estimation for different magnetic field strengths, compared with the results obtained from complete simulations including all effects by using realistic Gaussian bunch fields. The beam and device parameters correspond to Table I. The simulations were performed using the Virtual-IPM simulation tool [24]. The estimated distributions show agreement in the shape as well as the mean and variance. The additional tails of the calculated distributions emerge as a result of the electric field strength being independent of time in Eq. (1) through Eq. (4). Hence for electrons which are ionized before the bunch center the electric field is underestimated and it is overestimated for electrons which are ionized after the bunch center, leading to a corresponding under- and overestimation of the

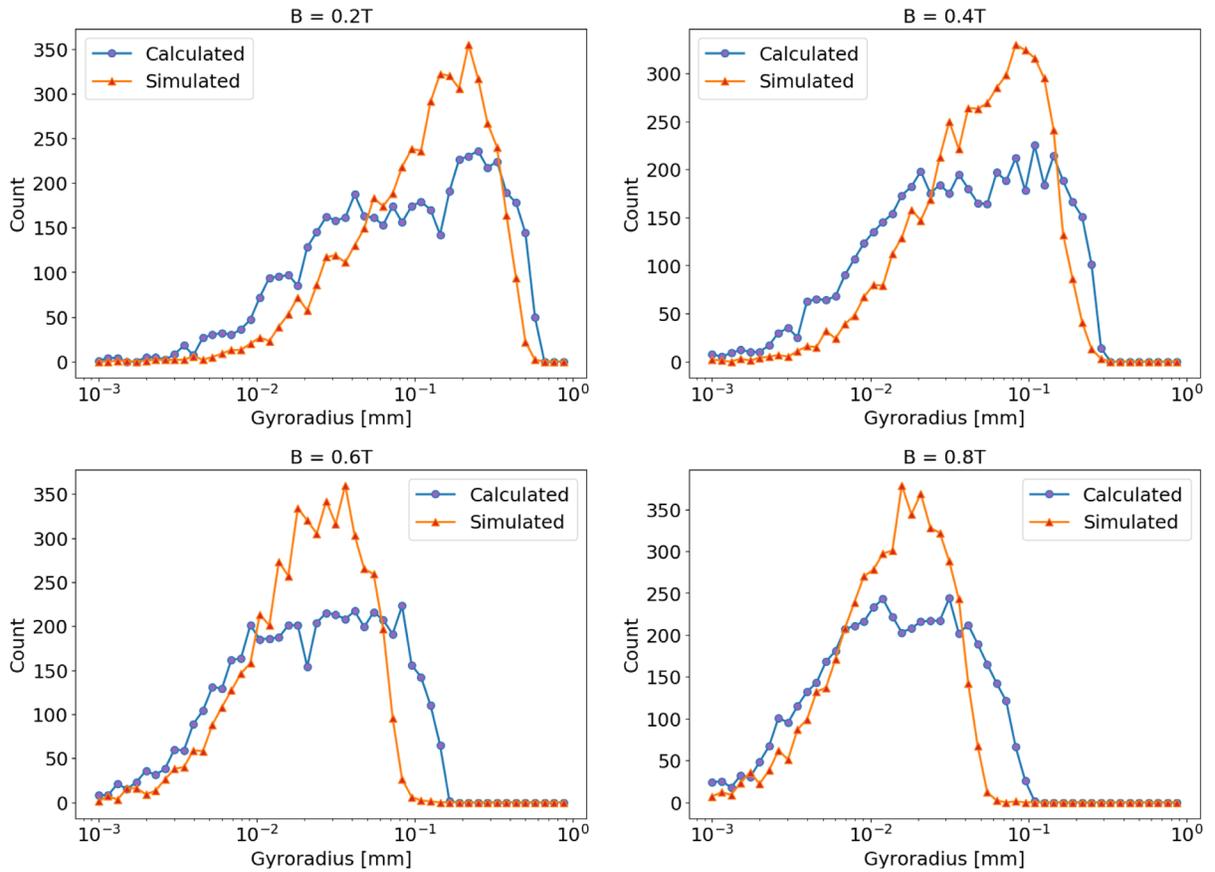


FIG. 9. Comparison of gyroradius distributions calculated from analytic equations with distributions obtained from complete simulations with realistic Gaussian bunch fields. Simulation parameters corresponding to Table I and the magnetic guiding field is varied.

resulting gyroradii which manifests itself in form of the tails of the distribution. The overall agreement of analytic prediction and simulation suggests the applicability of the derived equations. These equations, even though obtained for simplified bunch shape, can be used to further study the phenomenon of profile distortion as shown in the following section.

### C. Magnetic field tuning and scaling for realistic beams

Since IPMs find application in different circumstances including a variety of beam parameters it is useful to establish a relation between the beam parameters and the magnetic field required to suppress profile distortion. Such a relation allows for simple checks when designing new devices. Different criteria for a sufficient field strength are possible such as absolute thresholds ensuring that most of the electrons' gyroradii are below an acceptable threshold (e.g., requiring a certain fraction to be below some percentage of the bunch width or to be smaller than the detector resolution) or relative thresholds that ensure that the increase of measured profile standard deviation remains below some percent threshold. A previous study considered the gyroradius increase as a criterion [12] but since the relation between a particular gyroradius distribution and the resulting measured beam profile is not obvious we will focus on the standard deviation increase as a more direct measure for the further investigations. Note that due to the nature of profile distortion, which results in enlarged tails but also a heavier peak, already small values of standard deviation increase correspond to strongly non-Gaussian profile shapes as depicted in Fig. 6. We establish a fit of the

three most significant beam parameters relevant for profile distortion, the transverse and longitudinal bunch size and the bunch charge:

$$B_{\min,1\%} = \frac{N^a}{\sigma_z^b \sigma_t^c} \cdot d + \frac{f}{\sigma_t^e} \quad (8)$$

where  $a, b, c, d, e, f$  are fit parameters,  $\sigma_t$  is the standard deviation of the (symmetric) transverse beam distribution to be measured in units of millimeter,  $\sigma_z$  is the standard deviation of the bunch line density in units of nanosecond and  $N$  is the number of charges per bunch in units of  $1 \times 10^{12}$ .  $B_{\min,1\%}$  is the minimum magnetic field required to obtain  $\sigma_t^{\text{measured}}$  with at least 1% accuracy with respect to  $\sigma_t^{\text{beam}}$ . The first term in Eq. (8) corresponds to the space-charge interaction while the second term is attributed to the effect of the initial velocity distribution which, as a constant effect, has a relatively increasing effect on smaller beam profiles; hence the dependence on  $\sigma_t$ . The accuracy threshold was chosen to be a rather challenging 1% because this would be useful in investigation of emittance blow-up in modern colliders [25]. In a subsequent fit this threshold is converted to a parameter.

For verifying the relationship and to infer the parameters we use the analytic considerations from the previous section as well as simulations with three different bunch shapes: Gaussian, uniform and parabolic ellipsoid. For simplicity the bunch shapes are considered to be rotational symmetric around the beam axis ( $\sigma_t \equiv \sigma_x = \sigma_y$ ). The charge density distributions of these shapes are given by:

$$\begin{aligned} \text{Gaussian} & \propto \exp\left(-\frac{x^2 + y^2}{2\sigma_t^2} - \frac{z^2}{2\sigma_z^2}\right) \\ \text{Uniform} & \begin{cases} \propto \exp\left(-\frac{z^2}{2\sigma_z^2}\right) & , \text{ if } x^2 + y^2 \leq \sigma_t^2, \text{ with radius } \sigma_t \\ \equiv 0 & , \text{ otherwise} \end{cases} \\ \text{Parabolic Ellipsoid} & \begin{cases} \propto 1 - \frac{x^2 + y^2}{\sigma_t^2} - \frac{z^2}{\sigma_z^2} & , \text{ if } \frac{x^2 + y^2}{\sigma_t^2} + \frac{z^2}{\sigma_z^2} \leq 1, \text{ with semiaxes } \sigma_z, \sigma_t \\ \equiv 0 & , \text{ otherwise} \end{cases} \end{aligned} \quad (9)$$

We only consider relativistic beams since for low energy beams the ionization cross sections take on substantially different shapes [26], the generated electrons are exposed to the beam electric field for a longer period due to the reduced beam velocity and thus also effects from the longitudinal electric field start to play a role. We used relativistic  $\beta \geq 0.99$  when generating the fit data.

The relationship Eq. (8) is fitted with simulation data obtained with the Virtual-IPM simulation tool, for the parameters given in Table II. These parameters correspond to the Gaussian bunch shape. For the other bunch shapes the parameters are scaled where appropriate in order to

minimize the squared difference between initial beam distributions; for more details consider Appendix E. The resulting scaling factors are given in Table II as well. Table III shows the fit parameters obtained for the various bunch shapes as well as the quality of fit. They show agreement among each other and also with the results obtained for the analytic derivations. Applying the formula to the example case in Table I we obtain a minimum required field of 536 mT for suppressing the increase in standard deviation below 1%. A previous study, supported by a different simulation tool, estimated the minimum magnetic field required for distortionfree measurements to

TABLE II. Beam parameter ranges for simulation of the different bunch shapes. Scaling factors for minimizing the squared difference between the initial one-dimensional profiles of the different shapes are indicated as well. Each parameter follows a logarithmic distribution over 20 points resulting in a total of 8000 data points.

Bunch	-width $\sigma_t$	-length $\sigma_z$	-charge $N$
Gaussian	0.2 mm to 20 mm	0.3 ns to 300 ns	$1 \times 10^9$ to $1 \times 10^{13}$
Uniform	$\times 1.76$	$\times 1$	$\times 1$
Parabolic	$\times 2.44$	$\times 2.44$	$\times 1$
Ellipsoid			

TABLE III. Fit parameters for Eq. (8) for different bunch shapes as well as for the analytic considerations of Sec. II B. The quality of fit in form of mean squared error (MSE) is given as well.

	a	b	c	d	e	f	MSE [T <sup>2</sup> ]
Analytic	0.555	0.555	0.998	0.163	1.007	0.032	$3.65 \times 10^{-4}$
Uniform	0.609	0.605	1.013	0.110	0.986	0.049	$8.57 \times 10^{-4}$
Gaussian	0.613	0.590	1.082	0.105	0.967	0.038	$6.35 \times 10^{-4}$
Parabolic	0.623	0.595	1.027	0.103	0.974	0.046	$8.76 \times 10^{-4}$

be 1 T [13]. Using these two magnetic field strengths in our simulations we obtain an increase in standard deviation of 0.6 % and 0.08 % respectively which is in conformity with Eq. (8) and the results of the previous study. Applying the same formula to a set of beam parameters corresponding to the planned IPM on SIS-100 proton beam (29.9 GeV energy,  $2 \times 10^{13}$  charges,  $\sigma_t = 2.17$  mm,  $\sigma_z = 15$  ns), the minimum required field is estimated to 76 mT while the current design foresees a field of 84 mT [27].

A previous derivation of minimum required magnetic field strength that was done for related circumstances, however considering the gyroradius increase as an intermediate quantity instead of the final profile standard deviation, shows a similar dependence on the beam parameters [12].

In order to assess the effect of a chosen threshold for standard deviation increase  $\tau = \sigma_m \sigma_t^{-1} - 1$ , with measured (simulated) standard deviation  $\sigma_m$ , on the minimum required magnetic field we include this parameter as well in a subsequent fit of the form:

$$B_{\min} = \frac{N^a}{\sigma_z^b \sigma_t^c \tau^g} \cdot d + \frac{f}{\sigma_t^e \tau^h} \quad (10)$$

The results are similar to the ones obtained before and are shown in Table IV.

Another way to include the threshold in the formula is to model the remaining fit parameters as dependent on the threshold:

TABLE IV. Fit parameters for equation Eq. (10) for the Gaussian bunch shape. The quality of fit in form of mean squared error (MSE) is given as well.

a	b	c	d	e	f	g	h	MSE [T <sup>2</sup> ]
0.592	0.559	1.052	0.020	0.938	0.005	0.384	0.479	$3.79 \times 10^{-4}$

TABLE V. Fit parameters for equation Eq. (11) for the Gaussian bunch shape. The quality of fit in form of mean squared error (MSE) is given as well. For  $\tau = 1\%$  the results slightly deviate from the ones in Table III because the underlying simulation data was regenerated.

Threshold [%]	a	b	c	d	e	f	MSE [T <sup>2</sup> ]
1	0.626	0.584	1.065	0.103	0.974	0.044	$7.85 \times 10^{-4}$
2	0.605	0.571	1.064	0.093	0.942	0.031	$3.46 \times 10^{-4}$
3	0.596	0.553	1.073	0.082	0.927	0.025	$3.64 \times 10^{-4}$
4	0.572	0.536	1.061	0.077	0.903	0.021	$3.06 \times 10^{-4}$
5	0.571	0.538	1.039	0.071	0.927	0.019	$2.13 \times 10^{-4}$
6	0.574	0.538	1.047	0.064	0.898	0.018	$2.13 \times 10^{-4}$
7	0.571	0.554	1.051	0.059	0.878	0.017	$1.76 \times 10^{-4}$
8	0.565	0.545	1.029	0.056	0.884	0.016	$1.47 \times 10^{-4}$
9	0.551	0.544	0.990	0.055	0.907	0.014	$1.03 \times 10^{-4}$
10	0.547	0.553	1.000	0.051	0.891	0.014	$1.05 \times 10^{-4}$

$$B_{\min} = \frac{N^{a(\tau)}}{\sigma_z^{b(\tau)} \sigma_t^{c(\tau)}} \cdot d(\tau) + \frac{f(\tau)}{\sigma_t^{e(\tau)}}. \quad (11)$$

This formulation captures potential variations in the beam parameter dependence for varying thresholds  $\tau$ . The results however, as shown in Table V, confirm that the effect of varying threshold is mainly a scaling factor. These results are in agreement with the  $\tau^f$  scaling shown from Eq. (10) and Table IV.

### III. PROFILE CORRECTION

The first remedy against the space-charge profile distortion is to increase the applied magnetic guiding field, such that the gyroradii are bounded by user defined limits as shown by Eq. (8), Eq. (10), and Eq. (11) with corresponding parameters given by, respectively, Table III, Table IV, and Table V. These relations are fit for the broad parameter range in Table II. Most magnetic IPM designs fall into this category and further analysis is not required. However for some extreme scenarios very large magnetic fields are required [13], which are both expensive to obtain and occupy significant space in already cramped synchrotrons. As technology advances more applications for high energy and high brightness beams are to be expected, putting current magnetic IPM designs to test. For these scenarios, correction mechanisms to obtain a measure of actual profile from the distorted profile, have been studied.

Foremost of them was the study of a hardware electron “sieve,” which aims to filter electrons based on their gyroradius before they reach the acquisition system, and a numerical reconstruction on the resulting sieved profiles can be performed [14]. Though the study with simulations gave promising results, the sieve was found to be rather complex and thick structure difficult to integrate in machine vacuum. Another approach was parametric curve fitting of the distorted profile with analytic functions and correlation of the beam width with those parameters [13]. The fit result however suffered from too few available data points and hence a general relationship between the fit parameters and the beam profile could not be established.

Lately, approaches to record the inverse mapping between distorted profile to the original profile or the second moment of the original profile as a function of space charge parameters has been introduced either in form of look-up tables (LUT) [28] and supervised learning [20]. All the aforementioned correction methods rely on well understood and benchmarked IPM simulations. In this section, the problem of profile correction is generally introduced. Following that, several subapproaches in the supervised learning scheme are mentioned and an extension to a full profile reconstruction from distorted profiles with arbitrary initial profile shapes is shown.

### A. Problem description

In case of negligible space-charge interaction the distortion can be described via convolution of beam profile  $P_{\text{beam}}$  with a *point-spread function* (PSF) to obtain the measured profile  $P_{\text{measured}}$  [13]. This PSF depends on the initial velocities of electrons and can be assessed by means of differential ionization cross sections. The PSF itself describes the electron transport in the IPM. It corresponds to the probability that an electron is detected on a certain position displaced from its ionization position and can be obtained from considering the time that an electron spends above the various bins of the detector [14]. This PSF is independent of the position along the initial profile and hence the beam profile can be obtained by deconvolution.

If the PSF depends on the position along the profile—as in the case of space-charge influence on electron movement—a convolution cannot be used to describe the profile deformation anymore. As a more general transformation a matrix multiplication can be used instead:

$$P_{\text{measured},i} = M_{ij} \cdot P_{\text{beam},j} \quad (12)$$

where  $M_{ij}$  is the probability that an electron which was generated at position  $j$  is collected at position  $i$ . Comparing with convolution for the space-charge free case, the matrix  $M$  contains the PSF as columns, shifted across the rows with the PSF center at the diagonal ( $M_{ij} = \text{PSF}[i - j]$ ).

The transformation matrix depends on the beam parameters  $M_{ij} = M_{ij}(N_p, \sigma_x, \sigma_z)$  and for a given set of

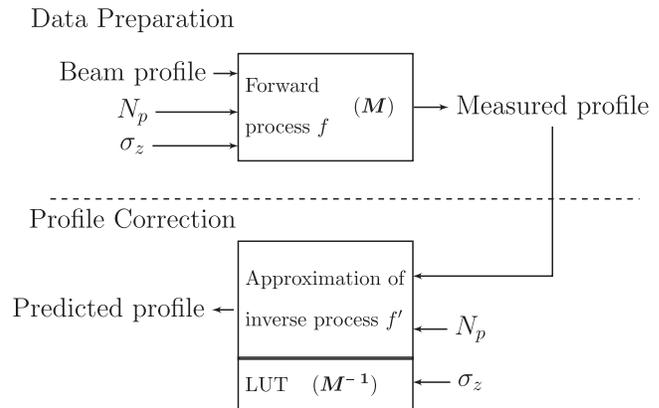


FIG. 10. The physical process is modeled by a simulation tool and the inverse process is approximated by the chosen machine learning algorithm or stored parameter dependent inverse mapping matrix.

parameters it can be obtained by simulations. Performing simulations over a grid of space-charge distortion parameters one can establish a look-up table of matrices  $M$  and their inverses  $M^{-1}$ . A fast iterative procedure to correct measured profiles using the simulated  $M^{-1}$  matrices is discussed in [28]. A newer approach of supervised machine learning is to deduce a set of rules for mapping measured profiles to their original counterparts. The idea is to infer a corresponding set of rules by providing distorted profiles alongside other relevant beam parameters and to map these to the original distribution of the beam. Figure 10 shows a schema of profile correction with the aid of simulation data.

### B. Data generation by IPM simulations

A recent joint effort between laboratories led to the development of a generic simulation tool called “Virtual-IPM” [24]. The Virtual-IPM tool has been used for simulating the movement of electrons inside the IPM, in presence of transient beam fields. Table VI shows the

TABLE VI. Parameters for the simulation. The bunch population, length, width, and height are varied, resulting in a total of 21 021 data samples. The parameter ranges roughly correspond to the LHC flat top.

Parameter	Values
Beam particle type	Protons
Energy/u	6.5 TeV
Bunch population	$1.1 \times 10^{11}$ ppb to $2.1 \times 10^{11}$ ppb
Bunch length ( $4\sigma_z$ )	0.9 ns to 1.2 ns
Bunch width ( $1\sigma_x$ )	270 $\mu\text{m}$ to 370 $\mu\text{m}$
Bunch height ( $1\sigma_y$ )	360 $\mu\text{m}$ to 600 $\mu\text{m}$
Electrode distance	85 mm
Applied voltage	4 kV
Magnetic field	0.2 T
Number of sim. electrons	1 000 000
Time step size	0.3125 ps

parameter ranges that are used for the simulations in order to span the relevant parameter space. The single bunches are modeled by three-dimensional Gaussian charge distributions. The electric field of bunches is computed via an analytic formula for a two-dimensional Gaussian charge distribution in the transverse plane [29] while the longitudinal dependency is taken into account by rescaling the field with the beam's line density. The longitudinal field component is neglected. This approximation is justified because the beam is highly relativistic and the longitudinal dimension of bunches is significantly larger than their transverse dimensions and therefore the electric field is mainly acting in the transverse plane. The advantage of the analytic formula, as compared to a numerical Poisson solver, is that it is much faster in computing the electric field and it does not suffer from discretization effects. Both the electric and the magnetic field of the beam are taken into account. The external electric and magnetic guiding fields are modeled to be uniform within the field cage. The initial velocities of electrons are generated according to a double differential cross section for a hydrogen target [22]. The passage of only a single bunch is simulated because the extraction times for electrons are only a few nanoseconds for the given electric guiding field while the bunch spacing is 25 ns.

The output of the simulations is summarized by histograms with 9.9 mm range and 55  $\mu\text{m}$  bin size, representing the electrons' positions at the moment of ionization, in the following referred to as *initial*, and at the moment of detection, in the following referred to as *final*. 55  $\mu\text{m}$  resolution corresponds to a hybrid-pixel detector type which has been successfully operated in the PS IPMs [18].

### C. Reconstruction of beam profile standard deviation

Equation (10) has the interesting property that for a given measured standard deviation of  $\sigma_m$  and a given magnetic field  $B$  the formula encodes the corresponding beam profile standard deviation  $\sigma_t$  which gives rise to the distortion via the threshold parameter  $\tau = \sigma_m \sigma_t^{-1} - 1$ . By rearranging Eq. (10) and inserting the definition of  $\tau$  the beam profile standard deviation emerges as the root of the function

$$f(\sigma_t) = \frac{d \cdot N^a}{\sigma_z^b \sigma_t^c (\sigma_t^{-1} \sigma_m - 1)^g} + \frac{f}{\sigma_t^e (\sigma_t^{-1} \sigma_m - 1)^h} - B = 0 \quad (13)$$

on the interval  $(0, \sigma_m)$ , which can be computed by means of numerical methods.

We test this method on the simulation data prepared according to Table VI by using bisection method [30] with 0.1  $\mu\text{m}$  tolerance. The resulting residuals have an overall mean and standard deviation of, respectively, 0.357  $\mu\text{m}$  and 3.21  $\mu\text{m}$ . The residuals plot Fig. 11 shows that the quality varies with beam profile standard deviation. Nevertheless most results are within 2% accuracy.

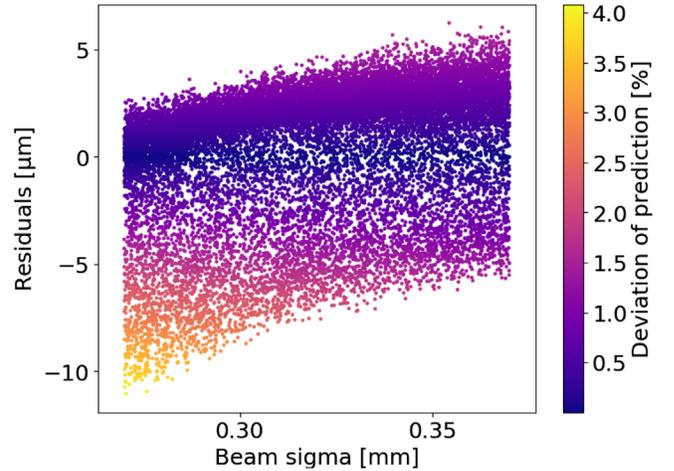


FIG. 11. Residuals for the beam profile standard deviation reconstruction based on Eq. (13) and tested with simulations corresponding to Table VI.

Previous attempts of reconstructing the initial beam profile standard deviation value from measured profiles using Machine Learning methods have shown a better performance while being fitted on a much smaller parameter range [20]. Equation (10) on the other hand has been fitted on a range which spans multiple orders of magnitude.

### D. Reconstruction of complete profiles with machine learning

Relevant for the presented problem are methods to perform supervised regression for predicting continuous variables [31]. In general a supervised machine learning (ML) model represents an algorithm  $f$ , a mapping from input  $x$  to output  $y_p$ , called *decision function*, that is specified by a set of parameters  $\theta \equiv \{\theta_i\}$ . The inputs and outputs are exactly opposite to the data preparation stage as shown in Fig. 8. While the structure of such an algorithm (e.g., the number of parameters  $|\theta|$ ) depends on its hyperparameters and is fixed, the goal is to tune  $\theta_i$  such that the corresponding function  $f(x|\theta)$  describes the output best. The quality of this description is typically assessed by a so-called *loss function*  $\mathbb{L}(x, y|\theta) \rightarrow \mathbb{R}^+$  which measures the deviation of the predictions  $y_p$  from the target output values  $y$ . Minimizing this loss function is used as a criterion for optimization of the model parameters  $\theta$ . In order to find a suitable ML algorithm for a given problem their performances are compared. For doing so the data set is split in three subsets corresponding to training, validation and testing. The training set is used to fit the particular ML algorithms in order to determine the optimal parameters  $\theta$ . The validation set is used for assessing the performance of an algorithm once it is fitted. Validating on a distinct data set prevents effects of over-fitting the algorithm to that particular training set and hence ensures generalization of the algorithm. The test similarly aims to prevent a tuning

bias toward the validation set, resulting from the multiple iterations corresponding to hyperparameter tuning. The test set is used to assess the final performance that is to be expected once a ML algorithm with a specific configuration has been chosen.

### 1. Full profile reconstruction

Previous works have studied the usage of machine learning models for establishing a relationship between measured profile and original beam profile standard deviation [19,20]. A very recent approach investigated reconstruction of the complete profile shape from measured profiles [21]. This approach used an additional global data transformation that is derived from the training data in order to converge during the fitting procedure. Here we present a novel approach that works only with per-profile normalization, providing a more consistent way of data preparation.

*Model architecture.*—In order to establish a mapping between measured profiles together with bunch length and bunch intensity data to the original beam profiles we recall Eq. (12). This equations states that the process of profile distortion can be described by a matrix multiplication  $M$  of the beam profile  $P_{\text{beam}}$ . In general this matrix is dependent on the beam parameters and hence the beam profile itself:  $M \rightarrow M(\sigma_z, N_p, P_{\text{beam}})$ . We make the assumption that the matrix is invertible and that the inverse matrix  $M^{-1}$  similarly depends on the measured profile:  $M^{-1} \rightarrow M^{-1}(\sigma_z, N_p, P_{\text{measured}})$ . Because the distortion arises from a symmetric process, the gyration of electrons, and the shift of final with respect to initial gyration center is negligible, as illustrated by Fig. 4 and discussed in Appendix B and Ref. [14], the columns of  $M$  are linearly independent and hence the invertibility criterion is fulfilled. The second assumption of  $M^{-1}$  being fully characterized by  $P_{\text{measured}}$  and  $\sigma_z, N_p$  is less obvious but it is vital for the following algorithm. In that sense the algorithm will only be successful if that condition is fulfilled. In the following we use simulation data to verify the validity of this assumption. Given  $M^{-1}(\sigma_z, N_p, P_{\text{measured}})$  the original profile can be obtained by means of a matrix multiplication  $M^{-1}$  with the measured profile  $P_{\text{measured}}$ . Here we identify the task of the neural network as the generation of the inverse matrix in dependence of the measured profile as well as the additional beam parameters bunch length and bunch intensity. Figure 12 sketches the procedure and shows the involvement of the machine learning model. Since the final prediction is obtained via a matrix multiplication (linear transformation) the neural network can be regarded as a generator model which produces linear models.

The architecture of the machine learning model and the neural network is depicted in Fig. 12. The neural network

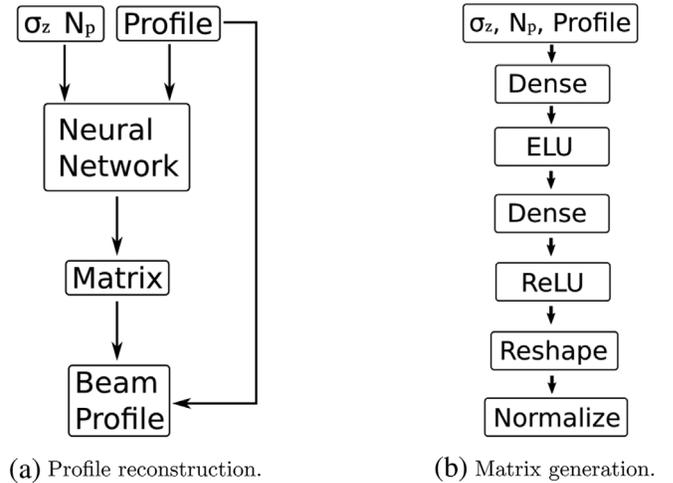


FIG. 12. Left: Sketch of the machine learning model which is used to perform the full profile reconstruction. The measured (distorted) profiles are used in two places, first to generate the transformation matrix and second for the subsequent dot product leading to the reconstructed beam profile. Right: Neural network architecture used for the matrix generation.

uses two dense layers, that is layers which connect all nodes from the previous layer with their own nodes. The layers use ELU and ReLU activation function respectively. ELU stands for exponential linear unit and is defined by  $\text{ELU}(x) = \{x \text{ if } x \geq 0 \text{ else } e^x - 1\}$ . ReLU stands for rectified linear unit and is defined by  $\text{ReLU}(x) = \max(0, x)$ . Hence, up to this stage, the used architecture resembles a feed-forward network with a single hidden layer, as depicted in Fig. 13, where the output layer represents the flat transformation matrix. ReLU activation is used on that output layer because the physical interpretation of matrix elements being fractions of transferred electrons allows only entries greater than or equal to zero. The flat output is then reshaped to a square matrix and each of its columns is normalized to unity due to the preservation of

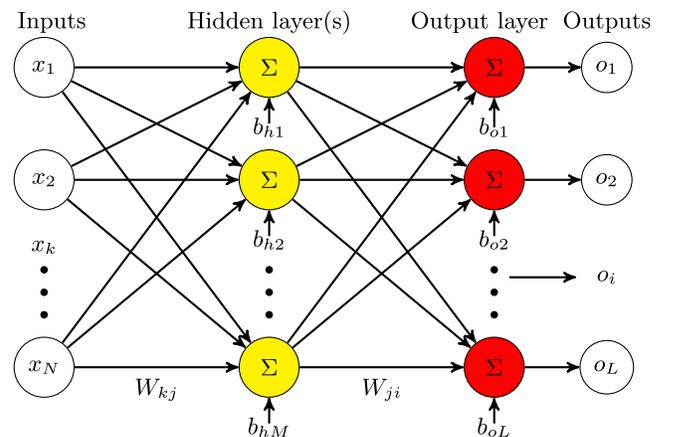


FIG. 13. Schematic of the architecture of a feed-forward multilayer perceptron.

TABLE VII. Hyperparameters that have been searched over and their values. Each hyperparameter set is randomly sampled from the given distribution. Batch size is log-sampled with base 2 and learning rate and number of nodes are log-sampled with base 10. A total of 1000 samples were scanned over.

Hyperparameter	Range
Batch size	1 to 2048
Learning rate	$1 \times 10^{-2}$ to $1 \times 10^{-6}$
Number of nodes	10 to 2288

the electron signal. We performed hyper-parameter search over the batch size, learning rate and the number of nodes in the first dense layer. The corresponding values are shown in Table VII. For identifying optimal configurations we monitored the final validation loss as well the ratio of initial to final validation loss in order to account for different starting losses. In a first step consensus about the batch size was found. With this optimal value fixed during a second and third iteration the optimal number of nodes and learning rate was identified respectively. The number of nodes however was found to have no significant influence on performance, if large enough, as shown in Fig. 14. The final configuration was batch size 236 (evenly divides the training set of 13 452 samples), learning rate  $8.882 \times 10^{-3}$  and number of nodes in the first dense layer equal to 42. We used He-uniform initializer [32] and mean squared error (MSE) as the loss function.

*Data preparation.*—Data preparation is an important aspect of any machine learning analysis and helps the fitting procedure to converge. Also it needs to be ensured that the way of data preparation is compatible with real application requirements. The previous study on full profile reconstruction [21] used per-profile normalization followed

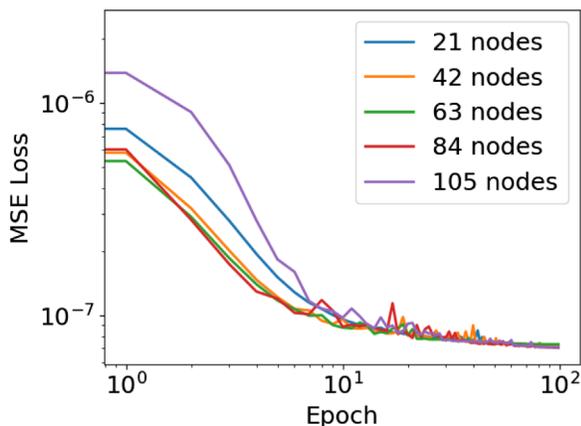


FIG. 14. Comparison of neural network performance for different widths of the first dense (hidden) layer. Other hyperparameters were fixed to initializer = He-uniform, batchsize = 236, learning rate =  $8.882 \times 10^{-3}$ .

by a per-feature normalization which was computed over the full training set. In the current study we focus on a more intuitive approach that only uses per-profile normalization. For data preparation we use the following steps: (1) Measured profiles are cropped to the range  $[-3.90 \text{ mm}, 3.90 \text{ mm}]$ . (2) Measured bunch length and number of charges per bunch are divided by, respectively, 1.2 ns and  $2.1 \times 10^{11}$ . (3) Measured profiles are normalized to unit integral. (4) Measured profile bins are set to zero for values smaller than 5% of the peak value.

The previous approach on the other hand computed a global valid range by considering a 5% peak-threshold region of the largest profile in the training set which was  $[-0.93 \text{ mm}, 0.93 \text{ mm}]$ . All profile bins outside that region were dropped and not used for the analysis. Hence, for comparison with the previous approach, we also apply this transformation in addition, by dropping data and predictions outside that valid range. However during training we use the full range available.

*Results.*—We used mean squared error (MSE) as a loss function and monitored the mean absolute error (MAE) as an additional metric. For Gaussian profiles we also compare the deviation of standard deviation of the target beam profiles and predicted profiles. Fitting converged after 25 epochs, yielding mean and standard deviation of  $\text{MSE} = 5.12 \times 10^{-7} \pm 2.49 \times 10^{-7}$ ,  $\text{MAE} = 6.44 \times 10^{-4} \pm 1.70 \times 10^{-4}$  on the test data set. The deviation of resulting profile standard deviation was  $\Delta\sigma_x/\sigma_x = 2.01\% \pm 1.00\%$  which is about one order of magnitude larger than what has been obtained from previous attempts with machine learning models directly targeting the profile standard deviation [20]. Table VIII shows the performance on the test set compared to the previously developed algorithm, the latter of which clearly performs better.

The neural network has been trained on Gaussian profiles only. However in real world applications the actual beam shapes might deviate from this ideal scenario. Hence an important requirement is that the reconstruction algorithm still works even if applied to different beam shapes. In order to verify this generalization we tested the performance on various non-Gaussian shapes, sampled from generalized Gaussian distributions and Q-Gaussian distributions. The charge distributions are given by the following equations.

TABLE VIII. Performance of old and new approach on the test data set, containing 4205 samples.

Loss	MSE [ $10^{-7}$ ]	MAE [ $10^{-4}$ ]	$\Delta\sigma/\sigma$ [%]	
Method	Meas.	Rec.	Rec.	
Old	$370 \pm 253$	$0.32 \pm 0.11$	$1.28 \pm 0.21$	$0.15 \pm 0.11$
New		$5.12 \pm 2.49$	$6.44 \pm 1.70$	$2.01 \pm 1.00$

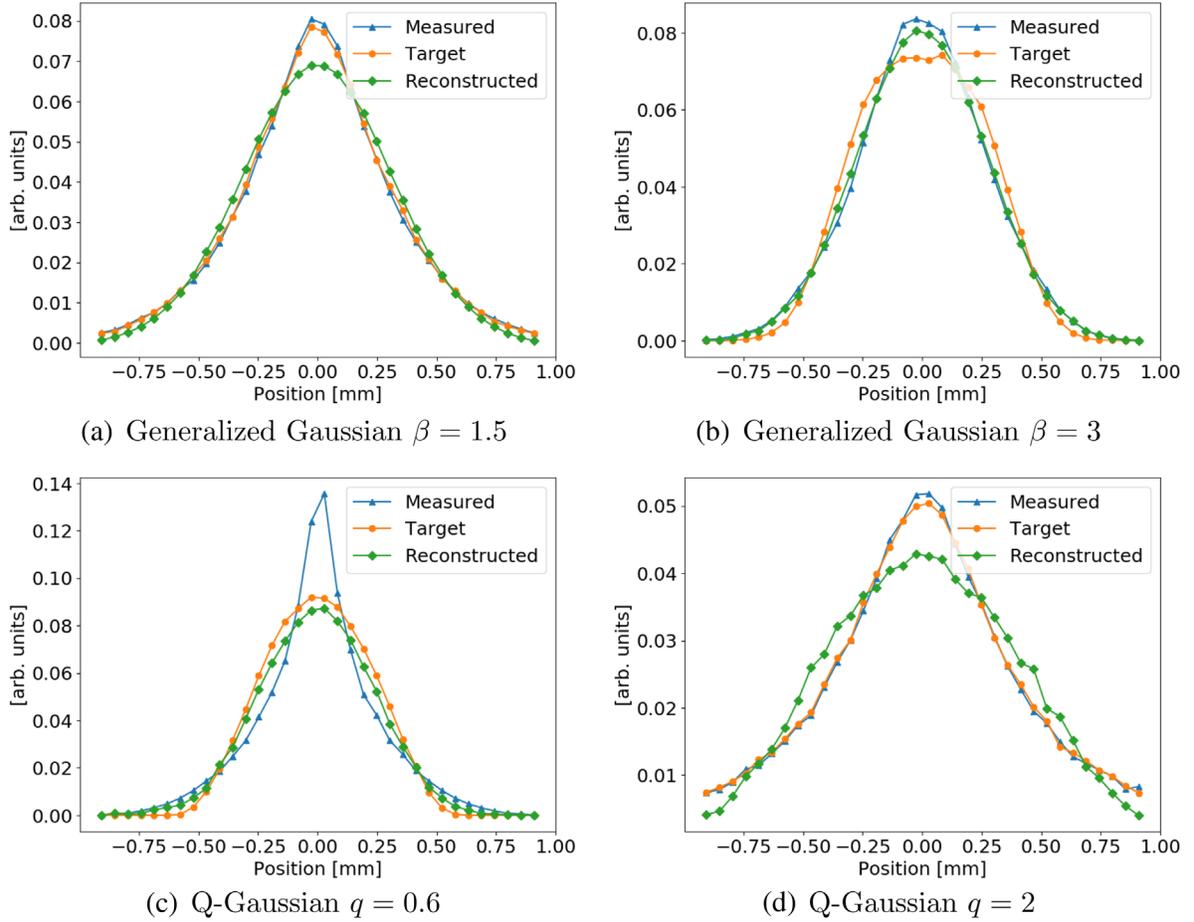


FIG. 15. Example profiles for non-Gaussian beam shapes together with reconstruction. “Measured” refers to the profile recorded by the detector (simulated in this case), “Target” refers to the initial beam distribution (the target of the reconstruction procedure), and “Reconstructed” refers to the result of the reconstruction procedure (which ideally should match the target). These examples are the ones with largest MSE error after reconstruction for each set. The Q-Gaussian profiles with  $q = 2$  (bottom right corner) are significantly wider than the training profiles.

Generalized Gaussian, used for instance in linac beam profile description [28]:

$$\frac{\beta}{2\alpha\Gamma(1/\beta)} \exp\left[-\left(\frac{|x-\mu|}{\alpha}\right)^\beta\right] \quad (14)$$

with shape parameters  $\alpha, \beta, \mu$ ;  $\beta = 2, \alpha = \sqrt{2}\sigma_x$  corresponds to a normal distribution.  $\Gamma$  denotes the gamma function [33].

Q-Gaussian, used in studies of tails of hadron beams [34]:

$$\frac{\sqrt{\beta}}{C_q} [1 - (1 - q)\beta x^2]^{-\frac{1}{q}} \quad (15)$$

with shape parameters  $\beta, q$ ;  $q = 1, \beta = \sigma_x^{-2}/2$  corresponds to a normal distribution.  $C_q$  is a normalization factor to provide unit integral.

The parameters of the non-Gaussian shapes correspond to the ranges shown in Table VI but are rescaled in order to match the Gaussian distribution in their limits. For the

generalized Gaussian  $\alpha_{x,y,z} = \sigma_{x,y,z}$  and for the Q-Gaussian  $\beta_{x,y,z} = \sigma_{x,y,z}^{-2}/2$ . Figure 15 shows an example profile for each profile shape together with the predicted reconstruction. Note that the Q-Gaussian profiles for  $q = 2$  are significantly wider than the training profiles and as a result show only very little distortion.

The overall performance, as shown in Table IX and Fig. 16, decreases about one order of magnitude as

TABLE IX. Test performance on various non-Gaussian beam shapes shown in Eq. (14) and Eq. (15), measured by mean square error (MSE) in units of  $10^{-6}$ . The mean and standard deviation of MSE performance is given.

Beam shape	Parameters	MSE <sub>meas</sub>	MSE <sub>rec</sub> <sup>old</sup>	MSE <sub>rec</sub> <sup>new</sup>
Gen. Gaussian	$\beta = 1.5$	$8.09 \pm 6.92$	$7.76 \pm 1.43$	$6.57 \pm 1.20$
	$\beta = 3$	$48.2 \pm 28.6$	$14.5 \pm 1.27$	$10.9 \pm 1.96$
Q-Gaussian	$q = 0.6$	$59.8 \pm 30.4$	$11.9 \pm 8.54$	$6.30 \pm 1.47$
	$q = 2$	$0.28 \pm 0.09$	$105 \pm 28.9$	$8.71 \pm 1.89$

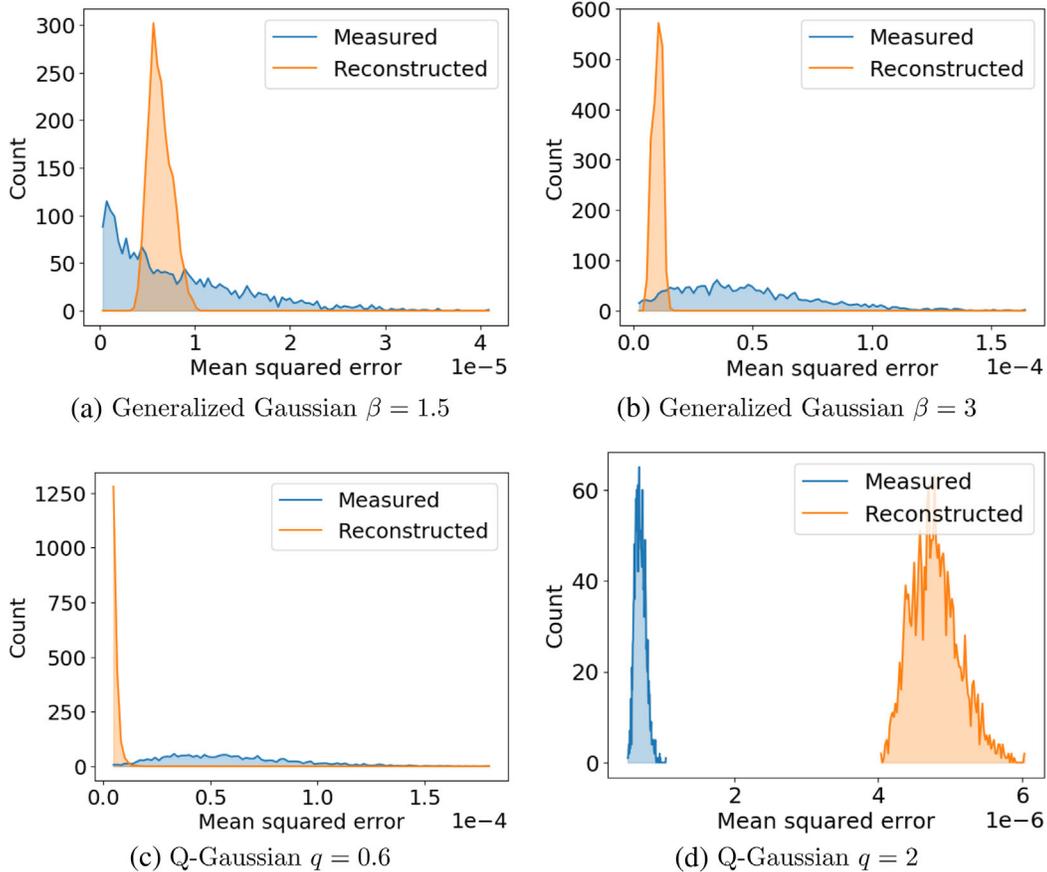


FIG. 16. Performance on non-Gaussian beam shapes. The Q-Gaussian profiles with  $q = 2$  (bottom right corner) are significantly wider than the training profiles.

compared to the Gaussian profiles. The new approach however shows improved performance when compared to the results of the previous approach. Because the wide  $q = 2$  Q-Gaussian profiles significantly exceed the input range of the algorithm the corresponding profile tail information is dropped and hence the performance degrades. This is however not an issue as one can always validate the range of measured profiles before feeding them into the algorithm. The obtained results on different profile shapes imply that the neural network identifies the underlying distortion mechanism rather than memorizing the specific profile shape which it has been trained on.

The results were obtained using various scientific PYTHON packages, including KERAS [35], MATPLOTLIB [36], NUMPY [37], PANDAS [38], SCIPY [39], TENSORFLOW [40], and VIRTUAL-IPM [24].

#### IV. SUMMARY

Distortions of measured beam profiles by electron-based magnetic ionization profile monitors is a relatively new issue affecting the successful operation of these devices at some present and future accelerators operated for high-brightness beams. We discussed the underlying effects and

their magnitude for this type of distortion in detail at the example of a possible LHC flat top configuration, supported by both analytic calculations as well as complete numerical simulations. Analytic descriptions for the simplified case of a uniform bunch shape are derived and benchmarking of these results with complete simulations was performed by using a well-established simulation tool. Simple formulas for computing the minimum magnetic field strength required to avoid space-charge induced profile distortion from relevant beam parameters are presented. They are validated against the derived analytic results for uniform beam as well as for complete simulations of various realistic beam shapes, including uniform, Gaussian, and parabolic shapes. Following that we addressed the problem of reconstruction of the real beam profile from distorted measurement. First an approach based on the previously developed formulas for minimum magnetic field is presented which allows us to infer the beam profile standard deviation from measured profile standard deviation on a broad range of beam parameters. The second approach addressed the problem of complete profile shape reconstruction and was realized by using a dedicated machine learning model involving neural network. The performance of this method was tested on

various realistic profile shapes, including Gaussian, Q-Gaussian and generalized Gaussian shapes. By fitting the model only with Gaussian profiles we showed that the approach generalizes as well to beam shapes significantly different than normal distributions.

### APPENDIX A: ESTIMATION OF MAXIMUM POLARIZATION DRIFT DISTANCE

The polarization drift velocity is described by the formula [23]

$$v(t) = \frac{1}{\omega} \frac{d}{dt} \left( \frac{E_{\perp}}{B} \right) \quad (\text{A1})$$

where  $\omega$  is the gyrofrequency of electrons and  $E_{\perp}$  is the electric field component perpendicular to the magnetic field  $B$ .

The maximum drift distance can be computed by

$$\delta = \int_0^{\infty} v(t) dt \quad (\text{A2})$$

where  $t = 0$  describes the time when the electron is longitudinally aligned with the bunch center, since for  $t < 0$  the electron drifts in the opposite direction. For estimating the bunch electric field we use the formula for the transverse electric field of a two-dimensional Gaussian bunch scaled with the longitudinal charge line density:

$$E_r(x, y, t) = \frac{Ne}{\epsilon_0} \frac{1}{\sqrt{2\pi^3} \sigma_z} \left[ 1 - \exp\left(-\frac{x^2 + y^2}{2\sigma_r^2}\right) \right] \exp\left(-\frac{\beta^2 c^2 t^2}{2\sigma_z^2}\right) \quad (\text{A3})$$

where  $N$  is the number of charges  $\sigma_r$  is the standard deviation of the transverse charge distribution and  $\sigma_z$  is the standard deviation of the longitudinal charge line density (in dimension of length);  $\beta$  is the relativistic beta factor of the particle beam.

We further neglect the time dependence of  $x$  and  $y$  because we consider  $\beta \approx 1$  together with nonrelativistic electrons and hence the change in beam electric field will be driven by the beam movement which is expressed through the explicit time dependence in Eq. (A3). The beam magnetic field is neglected as well since it is much smaller than the magnetic guiding field. Hence the electric field component  $E_{\perp}$  is determined to be perpendicular to the magnetic guiding field and is thus given by Eq. (A3).

It follows that the time derivative in Eq. (A1) is given by the partial time derivative of Eq. (A3). Integrating the resulting expression according to Eq. (A2) one obtains (dropping the sign):

$$\delta(x, y) = \frac{1}{\omega B} \frac{Ne}{\epsilon_0} \frac{1}{\sqrt{2\pi^3} \sigma_z} \left[ 1 - \exp\left(-\frac{x^2 + y^2}{2\sigma_r^2}\right) \right] \quad (\text{A4})$$

The maximum of the resulting function  $\delta(x)$  is given by:

$$\begin{aligned} \delta_{\max} &= \frac{1}{\omega B} \frac{N}{\sigma_r \sigma_z} \frac{e}{\epsilon_0} \frac{1}{\sqrt{2\pi^3} \sqrt{-\bar{w}}} \left[ 1 - \exp\left(\frac{\bar{w}}{2}\right) \right] \\ &\approx \frac{1}{\omega B} \frac{N}{\sigma_r \sigma_z} 5.18 \times 10^{-10} \end{aligned} \quad (\text{A5})$$

where  $\bar{w} = 2w_{-1}[-0.5 \exp(-0.5)] + 1$  and  $w_{-1}$  is the product log function [33].

Inserting the beam parameters from Table I and using  $\sigma_r = \sigma_x$  we obtain a maximum drift distance of  $\delta_{\max} = 852 \mu\text{m}$ .

### APPENDIX B: COMPENSATION OF GYRATION CENTER SHIFT

Equation (6) indicates the shift of the gyration center with respect to the undisturbed case (for which  $E = 0$ ). Using Eq. (5) this becomes:

$$x_c = \frac{\omega^2}{\Omega^2} \left( x_0 - \frac{v_{z0}}{\omega} \right) \quad (\text{B1})$$

with  $\omega = \bar{q}B$  the gyrofrequency for the undisturbed case  $E = 0$ .

For the derivation of Eq. (1)–Eq. (4) the bunch electric field was assumed constant, i.e.,  $E = \text{const}$ . In order to incorporate the time-dependence due to bunch movement we can consider a series of partial solutions  $(x_n(t), z_n(t))$  each of which is valid on an interval  $[t_n, t_{n+1}]$  for which the time dependence remains approximately constant. These partial solutions are of the form:

$$\begin{aligned} x_n(t) &= \frac{a_n}{\Omega_n} \sin(\Omega_n t) - \frac{b_n}{\Omega_n} \cos(\Omega_n t) + c_n \\ \dot{x}_n(t) &= a_n \cos(\Omega_n t) + b_n \sin(\Omega_n t) \\ z_n(t) &= -a_n \frac{\omega}{\Omega_n^2} \cos(\Omega_n t) - b_n \frac{\omega}{\Omega_n^2} \sin(\Omega_n t) + \frac{E_n}{B} c_n t + d_n \\ \dot{z}_n(t) &= a_n \frac{\omega}{\Omega_n} \sin(\Omega_n t) - b_n \frac{\omega}{\Omega_n} \cos(\Omega_n t) + \frac{E_n}{B} c_n \end{aligned} \quad (\text{B2})$$

where  $E_n \equiv E(t_n)$  and

$$\begin{aligned} a_0 &= v_{x0} \\ b_0 &= -\frac{\omega}{\Omega_0} \left( v_{z0} - \frac{E_0}{B} x_0 \right) \\ c_0 &= \frac{\omega^2}{\Omega_0^2} \left( x_0 - \frac{v_{z0}}{\omega} \right) \\ d_0 &= z_0 + \frac{\omega}{\Omega_0^2} v_{x0}. \end{aligned} \quad (\text{B3})$$

Note that  $c_0$  represents the (shifted) gyration center.

$$\begin{pmatrix} \frac{\sin(\Omega_n t_n)}{\Omega_n} & -\frac{\cos(\Omega_n t_n)}{\Omega_n} & 1 & 0 \\ \cos(\Omega_n t_n) & \sin(\Omega_n t_n) & 0 & 0 \\ -\frac{\omega}{\Omega_n^2} \cos(\Omega_n t_n) & -\frac{\omega}{\Omega_n^2} \sin(\Omega_n t_n) & \frac{E_n}{B} t_n & 1 \\ \frac{\omega}{\Omega_n} \sin(\Omega_n t_n) & -\frac{\omega}{\Omega_n} \cos(\Omega_n t_n) & \frac{E_n}{B} & 0 \end{pmatrix} \begin{pmatrix} a_n \\ b_n \\ c_n \\ d_n \end{pmatrix} = \begin{pmatrix} x_{n-1}(t_n) \\ \dot{x}_{n-1}(t_n) \\ z_{n-1}(t_n) \\ \dot{z}_{n-1}(t_n) \end{pmatrix} \quad (\text{B5})$$

The matrix on the left-hand side has determinant  $-\Omega\omega^{-1} \neq 0$  and hence a unique solution for the coefficients  $(a_n, b_n, c_n, d_n)$  exists. Inverting the matrix we obtain:

$$\begin{pmatrix} \frac{\Omega_n^2 - \omega^2}{\Omega_n} \sin(\Omega_n t_n) & \cos(\Omega_n t_n) & 0 & \frac{\omega}{\Omega_n} \sin(\Omega_n t_n) \\ \frac{\omega^2 - \Omega_n^2}{\Omega_n} \cos(\Omega_n t_n) & \sin(\Omega_n t_n) & 0 & -\frac{\omega}{\Omega_n} \cos(\Omega_n t_n) \\ \frac{\omega^2}{\Omega_n^2} & 0 & 0 & -\frac{\omega}{\Omega_n} \\ \frac{\omega \Omega_n^2 - \omega^3}{\Omega_n^2} t_n & \frac{\omega}{\Omega_n} & 1 & \frac{\omega^2 - \Omega_n^2}{\Omega_n^2} t_n \end{pmatrix}. \quad (\text{B6})$$

The parameter  $c_n$  encodes the gyration center shift and is given by the third line of the inverse matrix:

$$c_n = \frac{\omega^2}{\Omega_n^2} x_{n-1}(t_n) - \frac{\omega}{\Omega_n^2} \dot{z}_{n-1}(t_n). \quad (\text{B7})$$

Inserting Eq. (B2) into Eq. (B7) we obtain:

$$c_n = \frac{\Omega_{n-1}^2}{\Omega_n^2} c_{n-1} = c_0 \prod_{i=0}^{n-1} \frac{\Omega_i^2}{\Omega_{i+1}^2} = \frac{\Omega_0^2}{\Omega_n^2} c_0 = \frac{\omega^2}{\Omega_n^2} \left( x_0 - \frac{v_{z0}}{\omega} \right). \quad (\text{B8})$$

Since for  $n \rightarrow \infty$  we have  $\Omega_n \rightarrow \omega$  it follows that  $c_n \rightarrow x_0 - v_{z0}\omega^{-1}$  which is the undisturbed gyration center. Hence the initial shift, encoded in  $c_0$ , is compensated for  $t \rightarrow \infty$ . Actually this limit is already reached for  $E \rightarrow 0$  which is the case once the bunch has receded from the electron's position.

In order to link the solutions and their coefficients together we require continuity at the boundary of time intervals:

$$\begin{aligned} [x_n(t_n), \dot{x}_n(t_n), z_n(t_n), \dot{z}_n(t_n)] \\ \stackrel{!}{=} [x_{n-1}(t_n), \dot{x}_{n-1}(t_n), z_{n-1}(t_n), \dot{z}_{n-1}(t_n)]. \end{aligned} \quad (\text{B4})$$

Using Eq. (B2) we can write Eq. (B4) in matrix representation with respect to the coefficients:

### APPENDIX C: DERIVATION OF TIME-OF-FLIGHT FOR UNIFORM CHARGE DISTRIBUTION

We define  $a \equiv \frac{qE_a}{m}$ ,  $b \equiv \frac{q|E|}{m}$  and thus obtain:

$$\ddot{y}(t) = a - by \exp\left(-\frac{|t|}{\sigma_z}\right). \quad (\text{C1})$$

We solve the differential equation for  $t \geq 0$  (denoted as  $y^>$ ) since for  $t \leq 0$  (denoted as  $y^<$ ) we can then reuse this solution together with the transformation  $t \mapsto -t$ .

For  $t \geq 0$  the solution is:

$$\begin{aligned} y(t) &= \eta \sigma_z J_0[u(t)] G_{2,4}^{3,0} \left\{ \frac{u(t)}{2} \right\} \\ &\quad - \eta \sigma_z Y_0[u(t)] G_{2,0}^{1,3} \left\{ \frac{u(t)^2}{4} \right\} \\ &\quad + k_1 J_0[u(t)] + k_2 Y_0[u(t)] \end{aligned} \quad (\text{C2})$$

where  $\eta \equiv a\sigma_z\pi$ ,  $u(t) \equiv 2\sigma_z\sqrt{b} \exp(\frac{-t}{2\sigma_z})$ ,  $J_n, Y_n$  denote the Bessel functions of first and second kind respectively and  $G_{2,4}^{3,0}\{x\} \equiv G_{2,4}^{3,0}\{x, \frac{1}{2} |_{0,0,0,-\frac{1}{2}}\}$  and  $G_{2,0}^{1,3}\{x\} \equiv G_{2,0}^{1,3}\{x |_{0,0,0}\}$  denote Meijer G-functions [33].

By using the relations

$$\begin{aligned} \frac{d}{dx} G_{2,4}^{3,0}\{x\} &= -\frac{J_0(2\sqrt{x})}{x} \\ \frac{d}{dx} G_{2,0}^{1,3}\{x\} &= -\frac{2Y_0(2x)}{x} \end{aligned} \quad (\text{C3})$$

we confirm that the terms which contain the derivatives of the Meijer G-function cancel each other. We then obtain for  $\dot{y}(t)$ :

$$\dot{y}(t) = -\dot{u}(t) \left\{ \eta \sigma_z J_1[u(t)] G_{2,4}^{3,0} \left\{ \frac{u(t)}{2} \right\} - \eta \sigma_z Y_1[u(t)] G_{2,0}^{1,3} \left\{ \frac{u(t)^2}{4} \right\} + k_1 J_1[u(t)] + k_2 Y_1[u(t)] \right\} \quad (\text{C4})$$

Using the initial conditions  $y(0) \stackrel{!}{=} y_0$  and  $\dot{y}(0) \stackrel{!}{=} v_0$  we can derive the values for  $k_1, k_2$  (setting  $u_0 \equiv u(t_0)$ ):

$$\begin{pmatrix} J_0(u_0) & Y_0(u_0) \\ J_1(u_0) & Y_1(u_0) \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \end{pmatrix} = \begin{pmatrix} y_0 - \eta J_0(u_0) G_{2,4}^{3,0} \left\{ \frac{u_0}{2} \right\} + \eta Y_0(u_0) G_{2,0}^{1,3} \left\{ \frac{u_0^2}{4} \right\} \\ \frac{-v_{y0}}{u(0)} - \eta J_1(u_0) G_{2,4}^{3,0} \left\{ \frac{u_0}{2} \right\} + \eta Y_1(u_0) G_{2,0}^{1,3} \left\{ \frac{u_0^2}{4} \right\} \end{pmatrix} \quad (\text{C5})$$

For the  $t \leq 0$  case we apply  $t \mapsto -t$  and reuse the above solution. Note that the derivative  $\dot{u}(t)$  changes its sign due to the transformation. For  $t_0 < 0$  we obtain  $k_1^>, k_2^>$  from the continuity condition at  $t = 0$ :  $y^<(0) \stackrel{!}{=} y^>(0)$  and  $\dot{y}^<(0) \stackrel{!}{=} \dot{y}^>(0)$ .

#### APPENDIX D: METHODS FOR MINIMUM MAGNETIC FIELD COMPUTATION

For the computation of minimum magnetic field strength from analytic formulae we used the following methods. For each configuration (beam parameters + IPM parameters including magnetic field) 1000 particles were sampled from the initial beam distribution (sampling x- and y-position as well as the ionization time at  $z = 0$  since the guiding fields are assumed to be uniform and the resulting profile is integrated along  $z$ ). The initial momenta are sampled from double differential cross sections [22]. For each particle its time-of-flight (TOF) until leaving the bunch region, described by  $x(t)^2 + y(t)^2 \leq R^2$  and  $z(t) \leq (t - 4\sigma_z)\beta c$ , was computed as the minimum between the TOF in transverse direction, considering Eq. (1) and Eq. (C2), as well as in z-direction, using the bunch velocity and the electron velocity in z-direction via Eq. (2) and Eq. (4). We then use Eq. (3) and Eq. (4) in order to compute the final transverse velocity  $|v_{xz}(t_{\text{of}})| = \sqrt{v_x(t_{\text{of}})^2 + v_z(t_{\text{of}})^2}$  from which we calculate the gyroradius given the magnetic guiding field. We then consider gyration around the central point indicated by Eq. (1) and a random position of detection within the gyromotion range with probabilities proportional to the corresponding point-spread function (PSF) [14]. In order to increase the statistics, for each particle 100 final positions were sampled according to the corresponding PSF, all corresponding to the same initial position. This results in a total of 100 000 data points for computing the standard deviation of the beam profile and the measured profile respectively. Each configuration was scored according to the relative deviation of measured standard deviation with respect to standard deviation of the beam profile  $\lambda = (\sigma_m - \sigma_b)\sigma_b^{-1}$ , where  $\lambda$  denotes the score and  $\sigma_m, \sigma_b$  denote the measured and beam profile standard

deviation respectively. A particular magnetic field strength is considered sufficient if  $\lambda \leq 0.01$ . The magnetic field strength  $B$  is computed by finding the root of the function  $\lambda(B) - 0.01$  using the bisection method [30] with a tolerance of 1 mT.

#### APPENDIX E: SCALING FACTORS FOR DIFFERENT BUNCH SHAPES

Different bunch shapes are compared to a Gaussian distribution and their free parameters are adjusted such that the resulting distribution has a minimal least square difference to the corresponding Gaussian distribution:

$$\arg \min_{\eta} \int_{-\infty}^{+\infty} (\rho(x|\eta) - \mathcal{N}(0, \sigma))^p dx \quad (\text{E1})$$

where  $\rho(x|\eta)$  represents a particular bunch shape parametrized by  $\eta$ . Least squares difference ( $p = 2$ ) is used for the uniform bunch shape and  $p = 4$  is used for the parabolic ellipsoid shape for numerical stability. The L-BFGS-B solver from the scipy.optimize package [39] is used for minimizing Eq. (E1).

##### 1. Gaussian

The reference beam profile is a Gaussian distribution given by:

$$\rho_G(x) \propto \exp\left(-\frac{x^2}{\sigma_i^2}\right). \quad (\text{E2})$$

##### 2. Uniform

We consider a charge distribution that is uniform in the transverse shape and has a Gaussian dependency in the longitudinal direction:

$$\rho_U(x, y, z) \propto \begin{cases} \exp\left(-\frac{z^2}{2\sigma_z^2}\right) & , \text{ for } x^2 + y^2 \leq R^2 \\ 0 & , \text{ otherwise.} \end{cases} \quad (\text{E3})$$

Integration over  $z$  and  $y$  yields:

$$\rho_U(x|R) \propto \sqrt{R^2 - x^2}. \quad (\text{E4})$$

### 3. Parabolic ellipsoid

We consider a rotational symmetric ellipsoid with parabolic charge distribution:

$$\rho_P(x, y, z) \propto \begin{cases} 1 - \frac{x^2+y^2}{b} - \frac{z^2}{a^2} & , \text{ for } \frac{x^2+y^2}{b^2} + \frac{z^2}{a^2} \leq 1 \\ 0 & , \text{ otherwise.} \end{cases} \quad (\text{E5})$$

Integration over  $z$  and  $y$  yields:

$$\rho_P(x|\{a, b\}) \propto \frac{a}{b^3} (b^2 - x^2)^2. \quad (\text{E6})$$

- 
- [1] Simulation, design and operational experience with ionization profile monitors, ARIES workshop, GSI, May 2017, <http://indico.gsi.de/e/IPM17>.
- [2] J. Mießner, H.-J. Grabosch, R. Sternberger, M. Markert, K. Tiedtke, and A. Hofmann, An ionization profile monitor for the determination of the FLASH and PITZ beam parameter, in *Proceedings of the 2nd International Particle Accelerator Conference, San Sebastián, Spain* (EPS-AG, Spain, 2011).
- [3] P. Forck, Lecture notes on beam instrumentation, JUAS, 2018, [https://indico.cern.ch/event/683638/contributions/2801669/attachments/1563566/2540684/juas\\_script.pdf](https://indico.cern.ch/event/683638/contributions/2801669/attachments/1563566/2540684/juas_script.pdf).
- [4] K. Satou *et al.*, A novel field cage design for the CPS IPM and systematic errors in beam size and emittance, in *Proceedings of 9th International Particle Accelerator Conference (IPAC'18), Vancouver, BC, Canada, 2018* (JACoW Publishing, Geneva, Switzerland, 2018), pp. 2193–2196.
- [5] K. Satou and T. M. Mitsuhashi, Simulation and progress in ionization profile monitors for high intensity proton beam, in *Proceedings of International Beam Instrumentation Conference (IBIC'17), Grand Rapids, MI, USA, 2017* (JACoW, Geneva, Switzerland, 2018).
- [6] W. DeLuca, Beam detection using residual gas ionization, *IEEE Trans. Nucl. Sci.* **16**, 813 (1969).
- [7] R. E. Thern, Space-charge distortion in the Brookhaven ionization profile monitor, in *Proceedings of the 1987 Particle Accelerator Conference* (IEEE, Washington, DC, 1987).
- [8] W. Graves, Ph.D. thesis, Measurement of transverse emittance in the Fermilab booster, University of Wisconsin, 1994.
- [9] J. Amundson, J. Lackey, P. Spentzouris, G. Jungman, and L. Spentzouris, Calibration of the Fermilab Booster ionization profile monitor, *Phys. Rev. ST Accel. Beams* **6**, 102801 (2003).
- [10] F. Hornsta and M. Trump, A high resolution highly sensitive nonintercepting beam profile monitor for the Los Alamos Meson Facility (LAMPF), *Conf. Proc.* **C700928**, 1107 (1970).
- [11] M. Sapinski, W. Andreazza, B. Dehning, A. Guerrero, M. Patecki, and R. Versteegen, The first experience with LHC beam gas ionization monitor, in *Proceedings of IBIC 2012, Tsukuba, Japan, 2012* (JACoW, Geneva, Switzerland, 2013).
- [12] M. Patecki, M.Sc. thesis, Analysis of LHC beam gas ionization monitor data and simulation of the electron transport in the detector, 2013; Report No. CERN-THESIS-2013-155, 2013.
- [13] D. Vilsmeier, B. Dehning, and M. Sapinski, Investigation of the effect of beam space-charge on electrons in ionization profile monitors, in *Proceedings of HB 2014, East-Lansing, MI, USA, 2014* (JACoW, Geneva, Switzerland, 2015).
- [14] D. Vilsmeier, B.Sc. thesis, Profile distortion by beam space-charge in ionization profile monitors, 2015; Report No. CERN-THESIS-2015-035, 2015.
- [15] L. V. Keldysh, Ionization in the field of a strong electromagnetic wave, *JETP* **20**, 1307 (1965).
- [16] H. D. Zhang *et al.*, Development of a supersonic gas jet beam profile monitor, in *Proceedings of International Beam Instrumentation Conference (IBIC2015), Melbourne, Australia, 2015* (JACoW, Geneva, Switzerland, 2016), pp. 530–534.
- [17] C. C. Wilcox, B. Jones, A. Pertica, and R. E. Williamson, An investigation into the behaviour of residual gas ionisation profile monitors in the ISIS extracted beamline, in *Proceedings of International Beam Instrumentation Conference (IBIC'16), Barcelona, Spain, 2016* (JACoW, Geneva, Switzerland, 2017), pp. 808–811.
- [18] S. Lévassieur *et al.*, Time-resolved transverse beam profile measurements with a rest gas ionisation profile monitor based on hybrid pixel detectors, in *Proceedings of 9th International Particle Accelerator Conference (IPAC'18), Vancouver, Canada, 2018* (JACoW Publishing, Geneva, Switzerland, 2018), pp. 2361–2364.
- [19] R. Singh, M. Sapinski, and D. Vilsmeier, Simulation supported profile reconstruction with machine learning, in *Proceedings of International Beam Instrumentation Conference (IBIC'17), Grand Rapids, MI, USA, 2017* (JACoW, Geneva, Switzerland, 2018), pp. 350–354.
- [20] D. Vilsmeier, M. Sapinski, R. Singh, and J. W. Storey, Reconstructing space-charge distorted IPM profiles with machine learning algorithms, *J. Phys. Conf. Ser.* **1067**, 072003 (2018).
- [21] M. Sapinski, R. Singh, D. Vilsmeier, and J. W. Storey, Application of machine learning for the IPM-based profile reconstruction, in *Proceedings of 61st ICFA Advanced Beam Dynamics Workshop (HB'18), Daejeon, Korea, 2018* (JACoW Publishing, Geneva, Switzerland, 2018), pp. 410–415.
- [22] A. B. Voitkiv, N. Grün, and W. Scheid, Hydrogen and helium ionization by relativistic projectiles in collisions with small momentum transfer, *J. Phys. B* **32**, 3923 (1999).
- [23] R. Fitzpatrick, *Plasma Physics: An Introduction* (CRC Press, Taylor & Francis Group, 2014).

- [24] D. Vilsmeier, P. Forck, and M. Sapinski, A modular application for IPM simulations, in *Proceedings of International Beam Instrumentation Conference (IBIC'17), Grand Rapids, MI, USA, 2017* (JACoW, Geneva, Switzerland, 2018), pp. 355–358.
- [25] M. Kuhn *et al.*, Origins of transverse emittance blow-up during the LHC energy ramp, in *Proceedings of 5th International Particle Accelerator Conference IPAC2014, Dresden, Germany* (JACoW, Geneva, Switzerland, 2014), pp. 1021–1023.
- [26] M. E. Rudd, Y.-K. Kim, D. H. Madison, and J. W. Gallagher, Electron production in proton collisions: Total cross sections, *Rev. Mod. Phys.* **57**, 965 (1985).
- [27] M. Schwickert, O. Chorniy, T. Giacomini, P. Kowina, H. Reeg, T. Reichert, and R. Singh, Status of beam diagnostics for SIS100, in *Proceedings International Particle Accelerator Conference (IPAC'17), Copenhagen, Denmark, 2017* (JACoW, Geneva, Switzerland, 2017), pp. 156–159.
- [28] J. Egberts, Ph.D. thesis, IFMIF-LIPAc beam diagnostics: Profiling and loss monitoring systems, Université Paris Sud, Ecole Doctorale 534 MIPEGE, 2012.
- [29] M. Bassetti and G. A. Erskine, Closed expression for the electrical field of a two-dimensional Gaussian charge, Report No. CERN-ISR-TH/80-06, 1980 (unpublished).
- [30] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing* (Cambridge University Press, Cambridge, England, 2007).
- [31] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning* (Springer, New York, 2001).
- [32] K. He, X. Zhang, S. Ren, and J. Sun, Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification, *2015 IEEE International Conference on Computer Vision (ICCV) (2015)*, pp. 1026–1034, .
- [33] I. S. Gradshteyn and Iosif M. Ryzhik, *Table of Integrals, Series and Products*, 8th ed. (Academic Press, New York, 2014).
- [34] S. Papadopoulou, F. Antoniou, T. Argyropoulos, M. Fitterer, M. Hostettler, and Y. Papaphilippou, Modelling and measurements of bunch profiles at the LHC, *J. Phys. Conf. Ser.* **874**, 012008 (2017).
- [35] F. Chollet *et al.*, Keras, <https://keras.io>, 2015, version 2.2.4.
- [36] J. D. Hunter, Matplotlib: A 2D Graphics Environment, *Computing in Science Engineering* **9**, 90 (2007).
- [37] T. Oliphant, The NumPy Array: A Structure for Efficient Numerical Computation, *Computing in Science Engineering* **13**, 22 (2011).
- [38] W. McKinney, Data structures for statistical computing in Python, in *Proceedings of the 9th Python in Science Conference, Austin, Texas, 2010* (2010), pp. 51–56.
- [39] E. Jones, T. Oliphant, P. Peterson *et al.*, SciPy: Open source scientific tools for Python, <http://www.scipy.org/>, 2001, version 1.1.0.
- [40] M. Abadi *et al.*, TensorFlow: Large-scale machine learning on heterogeneous systems, 2015, Software available from [tensorflow.org](https://www.tensorflow.org), version 1.10.0.