

Estimation of longitudinal bunch characteristics in the LHC using Schottky-based diagnostics

Kacper Lasocha¹* and Diogo Alves²

CERN, CH-1211 Geneva 23, Switzerland

 (Received 10 December 2019; accepted 17 June 2020; published 26 June 2020)

The Large Hadron Collider (LHC) Schottky monitors have been designed to measure various parameters of relevance to beam quality, namely tune, momentum spread, and chromaticity. In this work, we present how this instrument can be used to estimate longitudinal bunch characteristics, such as longitudinal bunch profile or synchrotron frequency distribution. Under the assumption of bunched beams with no intrabunch coherent motion, we start by deriving the relation between the distribution of synchrotron amplitudes within the bunch population and the longitudinal bunch profile from probabilistic principles. Subsequently, we fit the cumulative power density of acquired Schottky spectra with the underlying distribution of synchrotron amplitudes. Finally, the result of this fit is used to reconstruct the bunch profile using the derived model. The results obtained with this method are verified by comparison with longitudinal profile measurements from the LHC wall current monitors.

DOI: [10.1103/PhysRevAccelBeams.23.062803](https://doi.org/10.1103/PhysRevAccelBeams.23.062803)

I. INTRODUCTION

The Large Hadron Collider (LHC) transverse Schottky system, whose main objective is to provide the beam operators with noninvasive bunch-by-bunch tune and chromaticity measurements, was commissioned in 2011 [1]. In the meantime, the system has undergone major upgrades in order to improve signal quality [2]. Although qualitatively its chromaticity estimates agree with other measurement techniques (as verified in dedicated experiments), the small residual quantitative discrepancies observed need to be understood [3]. Studies are therefore underway in order to better understand the obtained spectra. Due to its simpler physical interpretation, we have focused on the spectral region in the immediate vicinity of the 427725th revolution harmonic, where transverse particle motion plays no role. As by design the difference signal (rather than the sum signal) between two opposite slot-coupled waveguides is used, the inevitable common mode leakage allows us to use this spectral region for longitudinal studies.

As a result, a new application for the system has emerged, as a bunched-beam longitudinal profile monitor. In this work, we derive a relationship between the Schottky spectrum and the longitudinal bunch profile, under the

assumption of no intra-bunch coherent motion. A similar result, which uses the distribution function of the radio frequency (rf) bucket instead of the bunch profile, was presented in [4]. We also describe an alternative to the [5] method of obtaining the synchrotron frequency distribution. Finally, we provide a framework for the fast simulation of Schottky spectra, which can be used for further studies. The presented theory is supported by examples of LHC measurements and the obtained bunch profiles are compared with the wall current monitor (WCM) [6].

II. LONGITUDINAL BUNCH PROFILE

Within a bunch, the rf phase difference, $\Delta\phi_{\text{rf}}$, between a given particle and the synchronous particle, obeys equation [7] [Eq. (9.51)]:

$$\frac{d^2\Delta\phi_{\text{rf}}}{dt^2} + \Omega_{s_0}^2 \frac{\sin(\Delta\phi_{\text{rf}} + \phi_{\text{rf}_s}) - \sin(\phi_{\text{rf}_s})}{\cos(\phi_{\text{rf}_s})} = 0, \quad (1)$$

where Ω_{s_0} is the nominal synchrotron frequency and ϕ_{rf_s} is the rf phase of the synchronous particle. At constant beam energy, and for hadron machines such as the LHC where the energy loss per turn is small when compared to the maximum rf power, i.e., $\phi_{\text{rf}_s} \approx 0$, this equation reduces to the pendulum equation

$$\frac{d^2\Delta\phi_{\text{rf}}}{dt^2} + \Omega_{s_0}^2 \sin(\Delta\phi_{\text{rf}}) = 0. \quad (2)$$

For rf harmonic h , revolution frequency ω_0 and time amplitude (maximum time difference between a given particle and the synchronous particle) of synchrotron

*kacper.lasocha@cern.ch

Also at Institute of Physics, Jagiellonian University, Krakow, Poland.

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.

oscillations $\hat{\tau}$, we have that the particle's synchrotron frequency is given by:

$$\Omega_s = \frac{\pi}{2\mathcal{K}[\sin(\frac{h\omega_0\hat{\tau}}{2})]} \Omega_{s_0}, \quad (3)$$

where $h\omega_0\hat{\tau} = \widehat{\Delta\phi_{\text{rf}}}$ is the rf phase amplitude of synchrotron oscillations and $\mathcal{K}([0, 1]) \rightarrow [\pi/2, \infty]$ is the complete elliptic integral of the first kind [[8] p. 590]. This comes from the general theory of an arbitrary-amplitude pendulum [9].

From [10], we know that the time difference τ between a particle performing synchrotron motion and the synchronous particle can be approximated by a simple harmonic motion with amplitude-dependent Ω_s , i.e.,

$$\tau = \tau(\hat{\tau}, \phi_s) = \hat{\tau} \cos(\Omega_s t + \phi_s), \quad (4)$$

where $\Omega_s = \Omega_s(\hat{\tau})$. The validity of Eq. (4) can be confirmed by comparing it to the numerical solution of Eq. (1) and noting that $\Delta\phi_{\text{rf}} = h\omega_0\tau$. In the case of the LHC, Eq. (4) represents a good approximation for bunch lengths $\lesssim 80\%$ the size of the rf bucket, as illustrated in Fig. 1. The size of LHC rf bucket is 2.5 ns with the nominal 4-sigma bunch length less than 1.5 ns [11].

The longitudinal bunch profile can be interpreted as the probability distribution of τ . We shall denote this distribution by $\mathcal{B}(\tau)$. The assumption of no coherent intrabunch motion implies that the distribution of initial synchrotron phases, ϕ_s , is uniform and independent of the distribution of synchrotron amplitudes $\hat{\tau}$. Furthermore, under stationary conditions, the longitudinal bunch profile is independent of time. Therefore, the probability of finding any particle with time difference τ with respect to the synchronous particle can be written as a function of its amplitude of oscillation $\hat{\tau}$ only. We can then write:

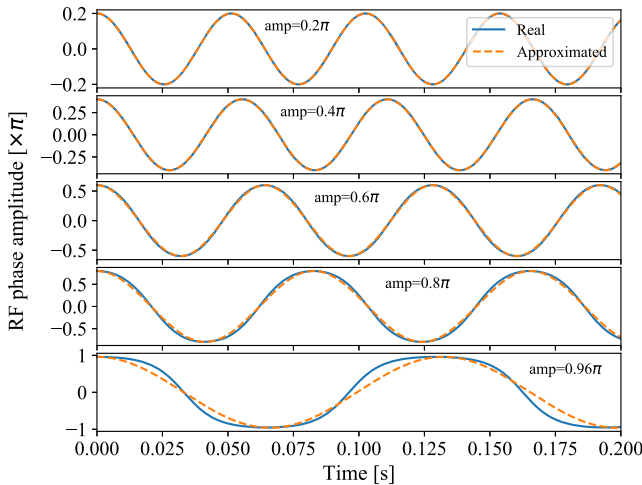


FIG. 1. Comparison between exact solution of Eq. (1) and approximation Eq. (4).

$$\mathcal{B}(\tau) = \int_0^\infty g_{\tau,\hat{\tau}}(\tau, \hat{\tau}) d\hat{\tau} = \int_{|\tau|}^\infty g_{\tau,\hat{\tau}}(\tau, \hat{\tau}) d\hat{\tau},$$

where $g_{\tau,\hat{\tau}}(\tau, \hat{\tau})$ is the joint probability density of a particle having amplitude $\hat{\tau}$ and time difference τ . The second equality comes from the fact, that $g_{\tau,\hat{\tau}}(\tau, \hat{\tau}) = 0$ for $|\tau| > \hat{\tau}$. The derivation of $g_{\tau,\hat{\tau}}(\tau, \hat{\tau})$ is presented in the Appendix and yields

$$\mathcal{B}(\tau) = \int_{|\tau|}^\infty \frac{g_{\hat{\tau}}(\hat{\tau})}{\pi\sqrt{\hat{\tau}^2 - \tau^2}} d\hat{\tau}. \quad (5)$$

This expression can be interpreted in an intuitive way. Let us first recall that $\mathcal{B}(\tau)$ is the probability density of having particles in the bunch with a time difference τ with respect to the synchronous particle. If we now think of $(\pi\sqrt{\hat{\tau}^2 - \tau^2})^{-1}$ as the probability density that a single particle, oscillating with amplitude $\hat{\tau}$, is found at a time distance τ from the synchronous particle, then all we have to do is to multiply it by the relative amount of particles, $g_{\hat{\tau}}(\hat{\tau})$, which oscillate with that same amplitude, and integrate over all possible amplitudes to finally obtain the longitudinal bunch profile, $\mathcal{B}(\tau)$. Figure 2 illustrates precisely the relation between $\mathcal{B}(\tau)$ and $g_{\hat{\tau}}(\hat{\tau})$ for different kinds of amplitude distributions. Note that even though $\tau = 0$ is the minimal point of every amplitude contribution, it is also the only point where all the amplitudes contribute.

Equation (5) allows us to calculate the longitudinal bunch profile knowing the distribution of synchrotron amplitudes. Conversely, if we are given the bunch profile, we can extract $g_{\hat{\tau}}$ by numerically solving an integral equation [12]. In addition, as the synchrotron amplitudes are related to the synchrotron frequencies by Eq. (3), knowing one of these distributions allows us to determine the other two, as shown in Fig. 3.

III. SCHOTTKY SPECTRUM

The Schottky spectrum is composed of a series of Bessel satellites (J_p) of finite width due to the presence of many particles with different synchrotron frequencies. The intensity signal due to a single-particle i , in the vicinity of the n -th revolution harmonic, can be written in the following form [10]

$$I_i(t) \propto \Re \left(\sum_{p=-\infty}^{\infty} j^p J_p(n\omega_0\hat{\tau}_i) e^{j(n\omega_0 t + p\Omega_{s_i} t + p\phi_{s_i})} \right). \quad (6)$$

Based on the above equation, we can deduce, that a single-particle power spectral density (PSD) is deterministic, and assuming that all macroscopic parameters ($n, \omega_0, \Omega_{s_0}$) are set, it depends only on particle's synchrotron amplitude $\hat{\tau}_i$. It is so, because synchrotron frequency can be derived from synchrotron amplitude using Eq. (3). We can see an example of such a spectrum on the upper plot of Fig. 4. Bessel

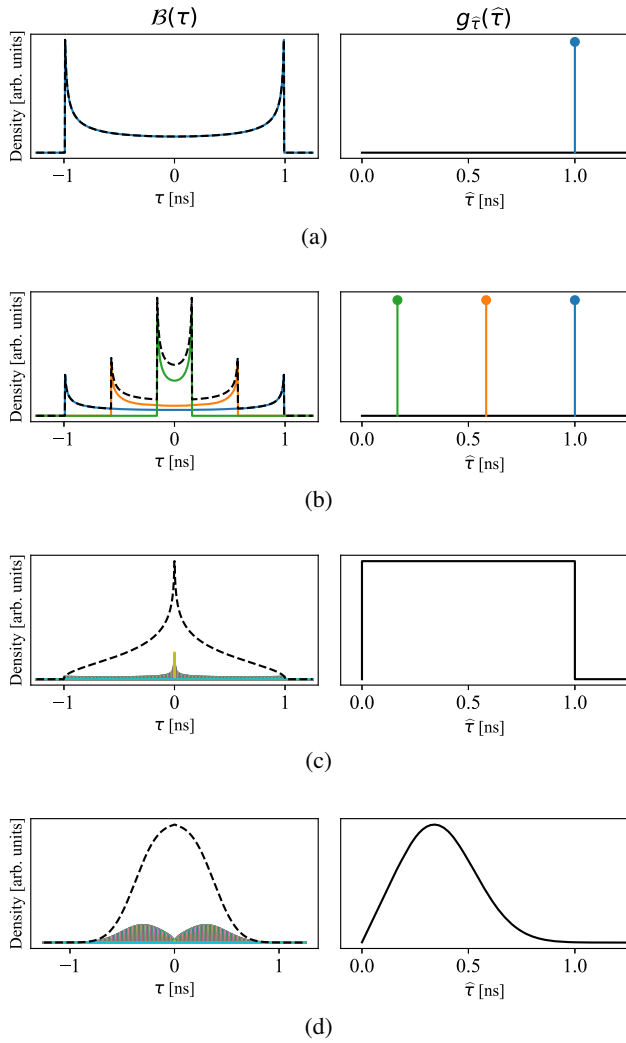


FIG. 2. Overview on possible synchrotron amplitude distributions and corresponding bunch profiles derived from Eq. (5). On the left side black dashed line denotes overall bunch shape, color solid lines are discrete amplitude contributions. On the right side distributions are given either by color Dirac delta functions or by solid black lines. For illustrative purposes, examples of discrete amplitude contributions are shown on the left plots even in the case of continuous amplitude distributions (a) All particles have the same synchrotron amplitude. (b) Particles form three amplitude groups. (c) Continuous uniform distribution of amplitudes. (d) Realistic non-uniform distribution of amplitudes.

satellites are evenly and symmetrically spaced, separated by the distance of Ω_{s_i} . The power of the satellite $\pm p$ is given by the value of $J_p^2(n\omega_0\hat{\tau}_i)$.

If we have many particles with different synchrotron amplitudes, the instantaneous Schottky spectrum does not look as simple anymore, see bottom plot of Fig. 4. Depending on the value of $n\omega_0\hat{\tau}_i$, each particle contributes in its own way and in different frequency ranges (the value of $J_p^2(n\omega_0\hat{\tau}_i)$ converges monotonically to zero with p for $p > n\omega_0\hat{\tau}_i$). Different values of Ω_{s_i} result in a broadening

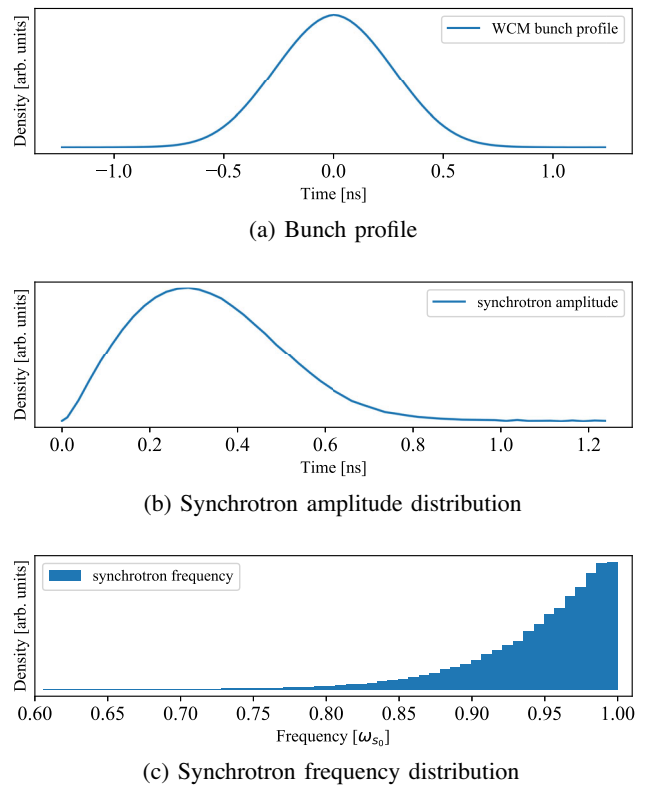


FIG. 3. Top: typical bunch profile at flattop, measured by a wall current monitor; Middle: synchrotron amplitude distribution derived from Eq. (5); Bottom: synchrotron frequency distribution derived from Eq. (3).

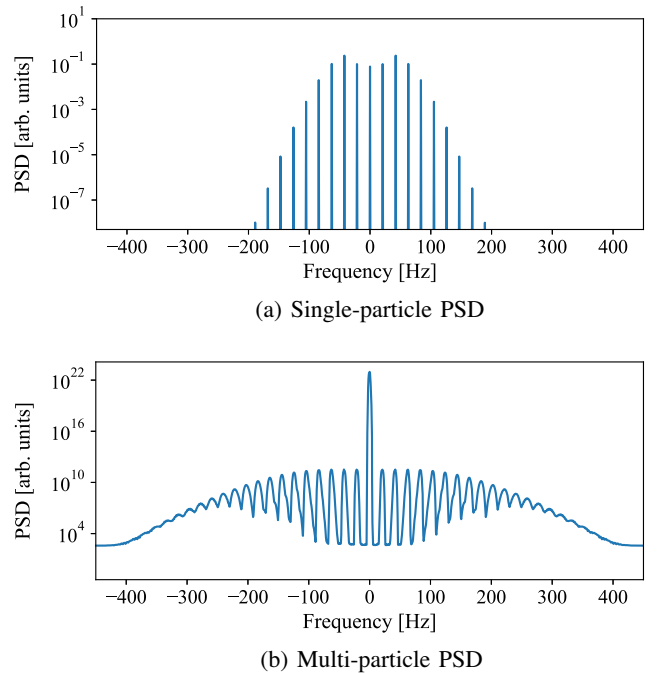


FIG. 4. Simulated longitudinal Schottky spectra of a single particle (top) and of an ensemble of 10^{11} particles (bottom). The zero frequency corresponds to an integer ($n = 427725$) harmonic of the revolution frequency.

and, for high indices of p , even overlapping of Bessel satellites. Due to the $p\phi_{s_i}$ phase term in Eq. (6), which vanishes for $p = 0$, contributions to the central satellite add up coherently, thus making its power density proportional to the square of the number of particles. For all other satellites, the instantaneous power density is nondeterministic, since it depends on the random synchrotron phases of all particles, and is proportional to the number of particles in the bunch.

Let us consider the PSD, $P(\omega)$ (where $\omega \neq n\omega_0$), of a pick-up signal $I(t) = \sum_i^N I_i(t)$, which is the sum of the individual contributions of N particles. As $I(t)$ is a wide-sense stationary process, the Wiener-Khinchin theorem [13] gives us

$$P(\omega) = \int_{-\infty}^{\infty} c(\tau) e^{-j\omega\tau} d\tau,$$

where $c(\tau) = \langle I(t)I^*(t-\tau) \rangle$ is the autocorrelation function of $I(t)$, $\langle \cdot \rangle$ denotes ensemble averaging and I^* denotes the complex conjugate of I . The previous equation can also be written in the following form

$$\begin{aligned} P(\omega) &= \int_{-\infty}^{\infty} \left\langle \left[\sum_{i=1}^N I_i(t) \right] \times \left[\sum_{i=1}^N I_i^*(t-\tau) \right] \right\rangle e^{-j\omega\tau} d\tau \\ &= \int_{-\infty}^{\infty} \left\langle \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N I_i(t) I_j^*(t-\tau) + \sum_{i=1}^N I_i(t) I_i^*(t-\tau) \right\rangle \\ &\quad \times e^{-j\omega\tau} d\tau. \end{aligned}$$

As $I(t)$ is a wide-sense stationary process, we do not need to specify time t . Since the synchrotron phases are uniformly and independently distributed, the first sum term of the expected value vanishes. In the second one the synchrotron phases cancel out, so we can write:

$$\begin{aligned} P(\omega) &= \int_{-\infty}^{\infty} \sum_{i=1}^N I_i(t) I_i^*(t-\tau) e^{-j\omega\tau} d\tau \\ &= \int_{-\infty}^{\infty} \sum_{i=1}^N \sum_{p=-\infty}^{\infty} J_p^2(n\omega_0\hat{\tau}_i) e^{j(n\omega_0\tau + p\Omega_{s_i}\tau)} e^{-j\omega\tau} d\tau \\ &= \sum_{i=1}^N \sum_{p=-\infty}^{\infty} J_p^2(n\omega_0\hat{\tau}_i) \delta(\omega - n\omega_0 - p\Omega_{s_i}). \end{aligned}$$

The above equation can be expressed in terms of discrete frequencies, written in matrix form as:

$$\underbrace{\begin{bmatrix} P_{\text{DFT}}(\omega_1, \hat{\tau}_1) & \cdots & P_{\text{DFT}}(\omega_1, \hat{\tau}_n) \\ P_{\text{DFT}}(\omega_2, \hat{\tau}_1) & \cdots & P_{\text{DFT}}(\omega_2, \hat{\tau}_n) \\ \vdots & \ddots & \vdots \\ P_{\text{DFT}}(\omega_m, \hat{\tau}_1) & \cdots & P_{\text{DFT}}(\omega_m, \hat{\tau}_n) \end{bmatrix}}_{\mathcal{M}} \cdot \underbrace{\begin{bmatrix} \tilde{g}(\hat{\tau}_1) \\ \tilde{g}(\hat{\tau}_2) \\ \vdots \\ \tilde{g}(\hat{\tau}_n) \end{bmatrix}}_{\mathcal{A}} = \underbrace{\begin{bmatrix} P_{\text{DFT}}(\omega_1) \\ P_{\text{DFT}}(\omega_2) \\ \vdots \\ P_{\text{DFT}}(\omega_m) \end{bmatrix}}_{\mathcal{S}}. \quad (9)$$

¹Every continuous distribution may be approximated with an arbitrary precision by a discrete distribution.

If we now denote the PSD of $I_i(t)$ as $P_i(\omega)$ we can write

$$P(\omega) = \sum_{i=1}^N P_i(\omega) \quad (7)$$

so the PSD is just the sum of single particle contributions.

IV. MATRIX FORMALISM

From the previous section we know that, in the absence of intrabunch coherent motion, the time averaged cumulative power spectrum of N particles is equal to the sum of the individual particle spectra. From Eq. (6) we know that differences in the individual particle spectra depend only on the particle's synchrotron amplitude, as synchrotron frequency can be expressed as a function of the amplitude [Eq. (3)]. Synchrotron phase does not influence the single particle's PSD. Therefore, the Schottky spectra are explicitly dependent on the distribution of synchrotron amplitudes.

Let us assume that we know the distribution $g(\hat{\tau})$ of synchrotron amplitudes among the particles. We may then calculate $P(\omega)$, the power at a given frequency, as:

$$P(\omega) = \int_0^{\infty} g(\hat{\tau}) P(\omega, \hat{\tau}) d\hat{\tau},$$

where $P(\omega, \hat{\tau})$ is the PSD at frequency ω of a particle with synchrotron amplitude $\hat{\tau}$. This can be seen as the continuous analogue of Eq. (7).

Experimentally, the power spectral density is estimated as the squared magnitude of the signal's discrete Fourier transform (DFT). We adopt the notation P_{DFT} for such an estimate. For a given frequency binning $\omega_1, \dots, \omega_m$, we shall have then

$$P_{\text{DFT}}(\omega_i) = \int_0^{\infty} g(\hat{\tau}) P_{\text{DFT}}(\omega_i, \hat{\tau}) d\hat{\tau}. \quad (8)$$

If we discretize $\hat{\tau}$ to a finite set of values $\hat{\tau}_j$,¹ Eq. (8) takes the form:

$$P_{\text{DFT}}(\omega_i) = \sum_j \tilde{g}(\hat{\tau}_j) P_{\text{DFT}}(\omega_i, \hat{\tau}_j),$$

where $\tilde{g}(\hat{\tau}_j)$ is the probability mass function of the discrete approximation of $g(\hat{\tau})$.

The columns of matrix \mathcal{M} correspond to the spectrum of a single particle with synchrotron amplitude $\hat{\tau}_i$. Vector \mathcal{A} represents the discrete approximation of the synchrotron amplitude density and vector \mathcal{S} is the DFT estimate of the total signal PSD, which can be compared with the experimentally obtained Schottky spectrum. As previously stated, considered frequencies ω_i must not cover frequency bins corresponding to the $p = 0$ satellite, as these add up coherently and Eq. (8) does not hold. One should also note that matrix \mathcal{M} depends on the nominal synchrotron frequency. Therefore, we shall use the notation $\mathcal{M}(\Omega_{s_0})$.

The expression of Eq. (9) is a very convenient tool for studying Schottky spectra. It enables us to simulate spectra, for different beam conditions and bunch shapes, without the need to perform time consuming Monte Carlo simulations. All we need to do is to calculate n single particle spectra with $\hat{\tau}$ ranging from $\hat{\tau}_1$ to $\hat{\tau}_n$.

V. BUNCH SHAPE CALCULATIONS

In order to estimate the longitudinal bunch shape, we gate our acquisition system on a single selected bunch and calculate its average discrete experimental spectrum \mathcal{S}_{exp} . Its length, which corresponds to the spectral resolution, depends on Schottky Monitor's sampling rate and desired sampling duration. As we aim for real-time measurements, a sampling duration of around 1 s was chosen, which in the case of LHC Schottky Monitor results in 32768 frequency bins and a spectral resolution of 0.69 Hz. It is high enough resolution to observe the inner structure of Bessel satellites, which is crucial for our further considerations. Extensive details on the acquisition architecture can be found in [2]. For all longitudinal studies' results presented in this paper, we limited ourselves to 2854 rows ($|\omega - n\omega_0| \in [20, 1000 \text{ Hz}]$). This frequency range depends on the nominal synchrotron frequency, has therefore to be different for LHC Injection and flat-top and will of course be different for other machines. As mentioned before, we need to exclude central bins, as well as corresponding rows from matrix $\mathcal{M}(\Omega_{s_0})$. Similarly, should other regions of the experimental spectrum be compromised by the presence of coherent peaks, they can also be omitted in the calculation of the $\mathcal{M}(\Omega_{s_0})$ matrix. The number of considered amplitudes $\hat{\tau}_i$ [number of columns of $\mathcal{M}(\Omega_{s_0})$] is a trade-off between the problem complexity and the discretization error of Eq. (9). From experience we determined that $n = 50$ is satisfactory for our objectives, however the impact of this parameter was not rigorously investigated.

The core of our approach is to minimize the cost function

$$C(\Omega_{s_0}, \mathcal{A}) = \|\log[\mathcal{M}(\Omega_{s_0}) \cdot \mathcal{A}] - \log[\mathcal{S}_{\text{exp}}]\|^2,$$

where the log functions are taken point-wise and $\|\cdot\|$ is the standard Euclidean norm. We use a logarithmic cost function here in order to be more sensitive to the low-magnitude spectral peripheries and more robust to noise, as

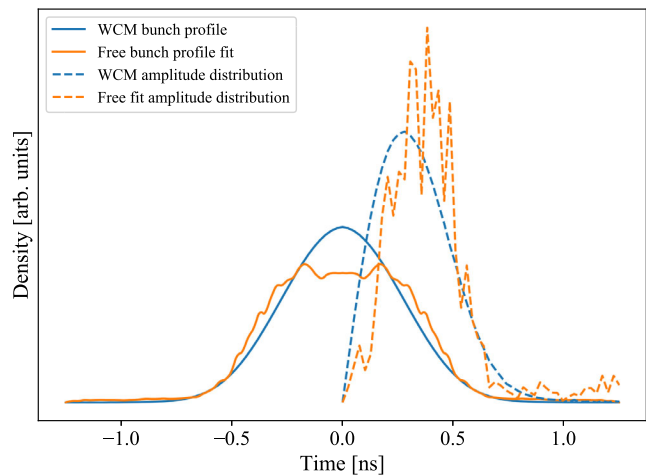


FIG. 5. WCM bunch profile and results of bunch profile fitting. Without putting any constraints on the synchrotron amplitude distribution we end up with an exotic bunch shape.

it was observed that the noise level in the spectrum is proportional to the spectrum value. Therefore, a usage of logarithm results in correct fitting to the measured data. Having found the optimal Ω_{s_0} and \mathcal{A} , we can estimate the longitudinal bunch shape and synchrotron frequency distribution using Eqs. (3) and (5) respectively.

It is certain that the experimental spectrum is susceptible to noise and finite time averaging effects. It may therefore happen that the pair $(\Omega_{s_0}, \mathcal{A})$ which minimizes $C(\Omega_{s_0}, \mathcal{A})$ is different from the true nominal synchrotron frequency and amplitude density. An example of such a situation is illustrated in Fig. 5, revealing exotic shapes for the estimated amplitude distribution and bunch shape, and where we have labeled this type of fit as “free fit”.

In order to overcome this feature, our proposed solution is based on the assumption that synchrotron amplitude densities follow a Rice distribution [14], which is the distribution of distances from the origin for samples taken from a circular 2D-normal distribution. It is determined by two parameters, standard deviation $\sigma_{\mathcal{N}_{2D}}$ and the modulus of mean $\nu_{\mathcal{N}_{2D}}$ of the mentioned 2D-normal distribution.² While a correlation between the Rice distribution and the distribution of synchrotron amplitudes is plausible at the moment of injection (the synchrotron amplitude is proportional to the distance from the origin in the longitudinal phase space), the question arises why it would remain valid after rf manipulations such as the ones performed during the energy ramp. Not knowing the answer to this question, we base our assumption on observations of bunch shapes measured by the WCM at different beam phases, and their corresponding synchrotron amplitude densities [calculated from Eq. (5)], which confirm this hypothesis. We present typical amplitude densities of LHC ion beams, together with the corresponding Rice

²Note that $\sigma_{\mathcal{N}_{2D}}$ and $\nu_{\mathcal{N}_{2D}}$ are generally not equal to the standard deviation and mean of the Rice distribution.

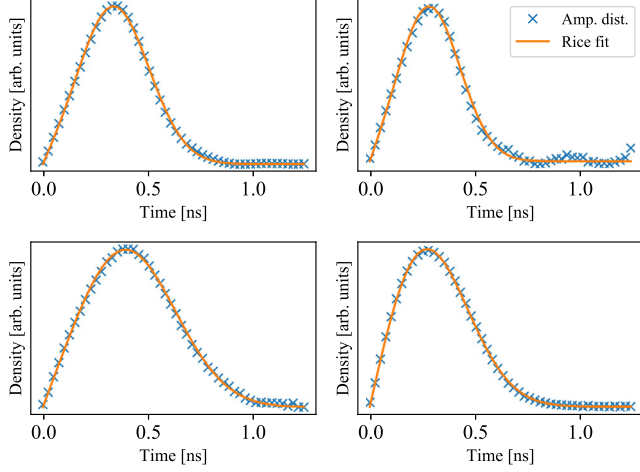


FIG. 6. Synchrotron amplitude distributions calculated for different beam modes (ramp, flat-top, injection, adjust) with comparable Rice distributions. Optimal Rice parameters were found as a result of curve fitting.

distribution fits in Fig. 6. Our assumption may be seen as a regularization, that is, introducing information which helps to solve an ill-posed problem by preventing solutions from wrongly compensating the errors. Furthermore, the number of fitting parameters is also dramatically reduced, from n parameter entries in vector \mathcal{A} ($n = 50$ in our implementation), to the 2 parameters: $\nu_{N_{2D}}$ and $\sigma_{N_{2D}}$.

Finding a solution to nonlinear problems is not always possible analytically. Therefore we decided to apply a differential evolution algorithm [15] implemented in a SCIPY library [16] in order to find the parameters which minimize the cost function $C(\Omega_{s_0}, \mathcal{A})$. Default algorithm settings were used, apart from $popsize = 60$ and $mutation = (0.5, 1.2)$.

Matrices $\mathcal{M}(\Omega_{s_0})$ were pre-calculated in order to reduce the computation time needed for the evaluation of the cost function. In order to determine the bunch shape and synchrotron frequency distribution, we fit 5 parameters in total. The first three have already been mentioned, these are μ and σ of the Rice distribution and the nominal synchrotron frequency. Additionally, we need to fit the scale c_1 , as the magnitude of the observed PSD may change, and finally, we need to take into consideration that some information may be masked by noise and we may actually only see the top part of the spectrum.³ It is worth noting, that the noise level is

³The scale parameter c_1 should be such, that values of $c_1 \mathcal{M}(\Omega_{s_0}) \cdot \mathcal{A}$ and \mathcal{S}_{exp} are comparable. As columns of $\mathcal{M}(\Omega_{s_0})$ have approximately the same sum and \mathcal{A} sums up to one, the scale parameter should be approximately equal to $\text{sum}\{\mathcal{S}_{\text{exp}}\} / \text{sum}\{\mathcal{M}_{.,i}(\Omega_{s_0})\}$, where $\text{sum}\{\mathcal{M}_{.,i}(\Omega_{s_0})\}$ is the sum of any single column. Due to effects of postprocessing (such as baseline subtraction), and errors of above mentioned approximations, we scan one order of magnitude range around the predicted value of c_1 . The offset parameter c_2 should be a fraction of $\max\{\mathcal{S}_{\text{exp}}\}$. In our implementation the whole range $[0, \max\{\mathcal{S}_{\text{exp}}\}]$ is scanned.

TABLE I. Parameter bounds used for SciPy differential evolution algorithm. Scale and noise bounds are not given, as they are strictly dependent on spectra processing.

$\sigma_{N_{2D}}$	$\nu_{N_{2D}}$	$f_{s_0} = \Omega_{s_0}/2\pi$
0–0.5 ns	$0-5\sigma_{N_{2D}}$	20–25 Hz (Flat-Top) 60–70 Hz (Injection)

subtracted from \mathcal{S}_{exp} during preprocessing. The final cost function including all these parameters takes the form

$$C(\Omega_{s_0}, \mathcal{A}, c_1, c_2) = |\log[\max\{c_1 \mathcal{M}(\Omega_{s_0}) \cdot \mathcal{A} - c_2, \mathcal{S}_{\text{min}}\}] - \log[\mathcal{S}_{\text{exp}}]|^2,$$

where \mathcal{S}_{min} is the minimal value of \mathcal{S}_{exp} . It turned out, that in the analysed spectra the noise parameter c_2 was observed to be negligible.

In order to increase convergence rate, parameters were bounded within the ranges specified in Table I, which are broad enough to include all possible physical solutions according to the machine configuration.

Figure 7 shows an example comparison between the experimental spectrum (in blue), the spectrum obtained after the “free fit” (in orange) and the spectrum obtained after the Rice fit (in green). It can be seen that both obtained spectra follow very well the overall experimental spectrum, including the fine details of the internal structure of the Bessel satellites. Finally, comparing the calculated profiles with the ones obtained from WCM measurements, confirms the accuracy of the proposed method. This is shown in Fig. 8 where a WCM profile measurement is compared with bunch profiles calculated from several spectra acquired from longitudinal spectra of both horizontal and vertical Schottky systems.

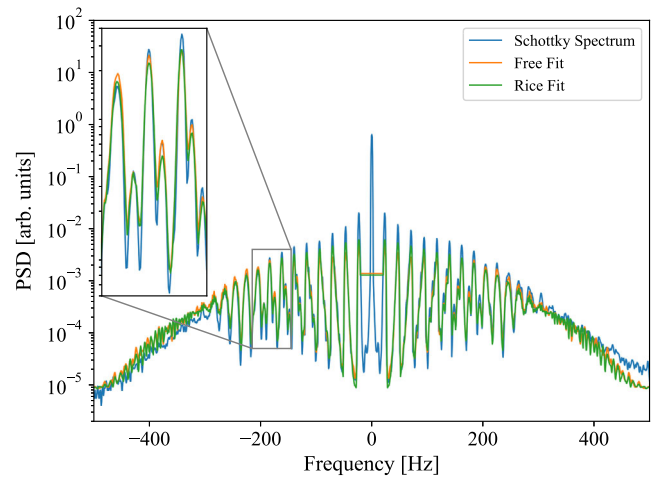


FIG. 7. Rice and free fit spectra compared with an experimental Schottky spectrum. There is no fit for the $p = 0$ satellite, as it adds up coherently and is therefore not described by Eq. (8).

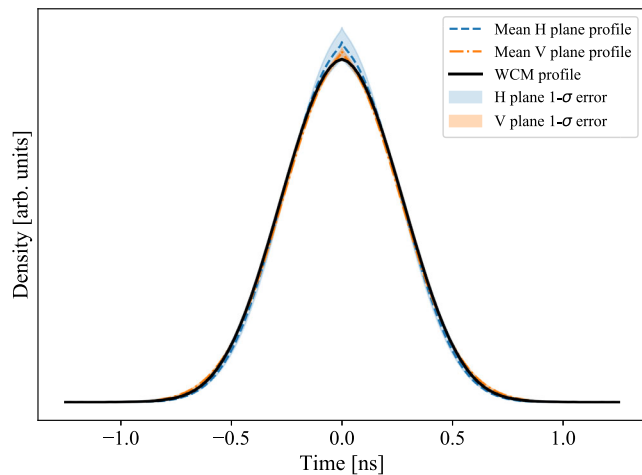


FIG. 8. Mean bunch shapes and $1 - \sigma$ error margins, estimated from the longitudinal spectra of horizontal and vertical LHC Schottky monitors in a time interval of 100 seconds (10 spectra per plane) around the WCM measurement.

VI. CONCLUSION

Within the scope of this paper, under the assumption of no intrabunch coherent motion, we have derived the relationship between three important bunch characteristics: longitudinal bunch profile, distribution of synchrotron amplitudes and distribution of synchrotron frequencies. Following, we have linked these characteristics to the Schottky spectrum in the form of matrix equation Eq. (9). Finally, we have presented a method of solving equation Eq. (9), that is estimating longitudinal bunch profile (as well as other mentioned characteristics) from the measured Schottky spectrum.

The results obtained have been verified in three stages. First, the experimental Schottky spectrum was compared to the one obtained from the optimization procedure showing that the Rice-based fit compares well with the observed PSD (and is also similar to the free fit). Secondly, bunch profiles calculated from vertical and horizontal monitors were observed to be self-consistent. Finally, comparing the calculated profiles with those obtained with the WCM confirms the accuracy of the proposed method.

The aim of this study is not to shift the main purpose of the LHC Schottky systems into a longitudinal profile measurement device, for which the WCMs provide a more direct diagnostic, but as an important step toward improving Schottky-based estimations in the LHC. The proposed procedure will now be adapted to transverse signals, which contain additional information on tune and chromaticity, and where the estimated bunch profiles are envisaged to be used also as a quality indicator for the derived quantities.

An additional advantage of the derived matrix formalism [Eq. (9)] is the fact that the fitting procedure does not require the use of all the points in the Schottky spectrum. For example, if certain spectral regions are affected by the presence of coherent peaks, they can be excluded from the

analysis by simply omitting the corresponding rows of matrix $\mathcal{M}(\Omega_{s_0})$. This linear formalism can also be used to predict the spectral response of Schottky diagnostics to various combinations of beam/machine parameters at different harmonics.

ACKNOWLEDGMENTS

The authors would like to acknowledge the stimulating discussions with O. R. Jones and T. Lefèvre, as well as the discussions and assistance of T. Argyropoulos, T. Levens, O. Marqvorsen, and M. Wendt.

APPENDIX: ANALYTICAL DERIVATION OF THE LONGITUDINAL BUNCH PROFILE

The aim of this Appendix is to derive Eq. (5). We shall start from

$$\mathcal{B}(\tau) = \int_0^\infty g_{\tau, \hat{\tau}}(\tau, \hat{\tau}) d\hat{\tau} = \int_{|\tau|}^\infty g_{\tau, \hat{\tau}}(\tau, \hat{\tau}) d\hat{\tau}.$$

Obtaining the explicit form of $g_{\tau, \hat{\tau}}(\tau, \hat{\tau})$ is not straightforward, as τ and $\hat{\tau}$ are not independent, but it can be derived from the joint distribution of initial synchrotron phases and amplitudes $g_{\phi_s, \hat{\tau}}$. We can write

$$g_{\phi_s, \hat{\tau}}(\phi_s, \hat{\tau}) = g_{\phi_s}(\phi_s) g_{\hat{\tau}}(\hat{\tau}) = \frac{g_{\hat{\tau}}(\hat{\tau})}{2\pi}, \quad (\text{A1})$$

as these random variables are independent and ϕ_s is uniformly distributed. In addition, let us define the transformation

$$u = (u_1, u_2) : (\phi_s, \hat{\tau}) \mapsto (\tau, \hat{\tau}),$$

where u_1 comes from Eq. (4) and u_2 is the identity function of $\hat{\tau}$:

$$\begin{aligned} u_1(\phi_s, \hat{\tau}) &= \hat{\tau} \cos(\Omega_s t + \phi_s), \\ u_2(\phi_s, \hat{\tau}) &= \hat{\tau}. \end{aligned}$$

Conversely, having the pair $(\tau, \hat{\tau})$, we can determine ϕ_s as one of the following:

$$\begin{aligned} \phi_{s_a}(\tau, \hat{\tau}) &= \arccos\left(\frac{\tau}{\hat{\tau}}\right), \\ \phi_{s_b}(\tau, \hat{\tau}) &= 2\pi - \arccos\left(\frac{\tau}{\hat{\tau}}\right). \end{aligned}$$

This gives us two possible inverse transforms

$$v_i^{-1} = (v_{1,i}, v_2) : (\tau, \hat{\tau}) \mapsto (\phi_s, \hat{\tau}),$$

where $i = a, b$ and functions $v_{1,i}, v_2$ are given by

$$\begin{aligned} v_{1,i}(\tau, \hat{\tau}) &= \phi_{s_i}(\tau, \hat{\tau}), \\ v_2(\tau, \hat{\tau}) &= \hat{\tau}. \end{aligned}$$

The relationship between the joint distributions of two sets of random variables related by known transformation functions is given in [[17] p. 201]. Using this, we obtain

$$g_{\tau, \hat{\tau}}(\tau, \hat{\tau}) = \sum_{i=a,b} g_{\phi_s, \hat{\tau}}(v_{1,i}(\tau, \hat{\tau}), v_2(\tau, \hat{\tau})) \left\| \begin{array}{cc} \frac{\partial v_{1,i}}{\partial \tau} & \frac{\partial v_2}{\partial \tau} \\ \frac{\partial v_{1,i}}{\partial \hat{\tau}} & \frac{\partial v_2}{\partial \hat{\tau}} \end{array} \right\|.$$

where by $\|\cdot\|$ we denote the absolute value of the determinant of a matrix. We calculate the Jacobian in the equation above by noting, that $\frac{\partial v_2}{\partial \tau} = \frac{\partial \hat{\tau}}{\partial \tau} = 0$ and $\frac{\partial v_2}{\partial \hat{\tau}} = \frac{\partial \hat{\tau}}{\partial \hat{\tau}} = 1$, so

$$\left| \begin{array}{cc} \frac{\partial v_{1,i}}{\partial \tau} & \frac{\partial v_2}{\partial \tau} \\ \frac{\partial v_{1,i}}{\partial \hat{\tau}} & \frac{\partial v_2}{\partial \hat{\tau}} \end{array} \right| = \left| \begin{array}{cc} \frac{\partial v_{1,i}}{\partial \tau} & 0 \\ \frac{\partial v_{1,i}}{\partial \hat{\tau}} & 1 \end{array} \right| = \frac{\partial v_{1,i}}{\partial \tau} = \pm \frac{1}{\sqrt{\hat{\tau}^2 - \tau^2}}.$$

We have then that

$$\begin{aligned} g_{\tau, \hat{\tau}}(\tau, \hat{\tau}) &= \sum_{i=a,b} \frac{g_{\phi_s, \hat{\tau}}(v_{1,i}(\tau, \hat{\tau}), v_2(\tau, \hat{\tau}))}{\sqrt{\hat{\tau}^2 - \tau^2}} \\ &= \sum_{i=a,b} \frac{g_{\phi_s}(v_{1,i}(\tau, \hat{\tau})) g_{\hat{\tau}}(v_2(\tau, \hat{\tau}))}{\sqrt{\hat{\tau}^2 - \tau^2}} \\ &= \sum_{i=a,b} \frac{g_{\hat{\tau}}(v_2(\tau, \hat{\tau}))}{2\pi\sqrt{\hat{\tau}^2 - \tau^2}} = \frac{g_{\hat{\tau}}(v_2(\tau, \hat{\tau}))}{\pi\sqrt{\hat{\tau}^2 - \tau^2}} \\ &= \frac{g_{\hat{\tau}}(\hat{\tau})}{\pi\sqrt{\hat{\tau}^2 - \tau^2}}, \end{aligned}$$

where we have separated $g_{\phi_s, \hat{\tau}}$ into the product of two distributions using Eq. (A1). It enables us to finally write

$$\mathcal{B}(\tau) = \int_{|\tau|}^{\infty} \frac{g_{\hat{\tau}}(\hat{\tau})}{\pi\sqrt{\hat{\tau}^2 - \tau^2}} d\hat{\tau}.$$

-
- [1] R. Pasquinelli and A. Jansson, Microwave Schottky diagnostic systems for the Fermilab Tevatron, recycler, and CERN large hadron collider, *Phys. Rev. Accel. Beams* **14**, 072803 (2011).
 [2] M. Betz, O. Jones, T. Lefevre, and M. Wendt, Bunched-beam Schottky monitoring in the LHC, *Nucl. Instrum. Methods Phys. Res., Sect. A* **874**, 113 (2017).

- [3] T. Tydecks, D. Alves, T. Levens, M. Wendt, and J. Wenninger, Status of the LHC Schottky Monitors, in *Proc. 9th International Particle Accelerator Conference (IPAC'18), Vancouver, BC, Canada, April 29-May 4, 2018*, International Particle Accelerator Conference No. 9 (JACoW Publishing, Geneva, Switzerland, 2018), pp. 247–249, <https://doi.org/10.18429/JACoW-IPAC2018-MOPMF058>.
 [4] V. Balbekov and S. Nagaitsev, Longitudinal schottky spectra of bunched beams, in *Proceedings of the 9th European Particle Accelerator Conference, Lucerne, 2004* (EPS-AG, Lucerne, 2004), EPAC-2004-MOPLT109 [<http://accelconf.web.cern.ch/AccelConf/e04/>].
 [5] E. Shaposhnikova, Longitudinal peak detected Schottky spectrum, CERN Technical Report No. CERN-BE-2009-010, 2009.
 [6] G. Papotti, F. Follin, T. Bohl, and U. Wehrle, Longitudinal beam measurements at the LHC: The LHC beam quality monitor, Particle accelerator, Proceedings, 2nd International Conference, IPAC 2011, San Sebastian, Spain, September 4-9, 2011, *Conf. Proc.* C110904, 1852 (2011).
 [7] H. Wiedemann, *Particle Accelerator Physics*, 4th ed. (Springer, Berlin, 2015).
 [8] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9th dover printing, 10th gpo printing ed. (Dover, New York, 1964).
 [9] K. Ochs, A comprehensive analytical solution of the nonlinear pendulum, *Eur. J. Phys.* **32**, 479 (2011).
 [10] D. Boussard, Schottky noise and beam transfer function diagnostics, *CERN Accelerator School: 5th Advanced Accelerator Physics Course* (1995), p. 749, <https://doi.org/10.5170/CERN-1995-006.749>.
 [11] O. S. Brning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report, CERN Yellow Reports: Monographs* (CERN, Geneva, 2004).
 [12] F. W. Stallmann, Numerical solution of integral equations, *Numer. Math.* **15**, 297 (1970).
 [13] D. Ricker, *Echo Signal Processing*, 1st ed. (Springer, New York, 2003).
 [14] A. Abdi, C. Tepedelenlioglu, M. Kaveh, and G. Giannakis, On the estimation of the k parameter for the Rice fading distribution, *IEEE Commun. Lett.* **5**, 92 (2001).
 [15] R. Storn and K. Price, Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces, *J. Global Optim.* **11**, 341 (1997).
 [16] E. Jones, T. Oliphant, P. Peterson *et al.*, SciPy: Open source scientific tools for Python (2001–), Version 1.1.0, [Online; accessed].
 [17] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*, 4th ed. (McGraw Hill, Boston, 2002).