Benchmarking coherent transition radiation spectroscopy as a tool for high-resolution bunch shape reconstruction at free-electron lasers

Bernhard Schmidt, Nils Maris Lockmann, Peter Schmüser, and Stephan Wesch
Deutsches Elektronen-Synchrotron DESY, Notkestrasse 85, 22607 Hamburg, Germany

(Received 11 April 2020; accepted 27 May 2020; published 16 June 2020)

We present a systematic comparison of two complementary methods for determining the longitudinal charge density profile of the compressed electron bunches in the soft x-ray free-electron laser FLASH: a frequency-domain technique—coherent transition radiation (CTR) spectroscopy—and a time-domain technique—streaking of the electron beam with a transversely deflecting microwave structure (TDS). While the direct time-profile measurement with a TDS is a well-established method invented at SLAC, our group has pioneered high-resolution bunch shape analysis based on coherent radiation spectroscopy. We have developed a broadband spectrometer covering the wavelength range from $5 \mu m$ to $433 \mu m$ with two sets of remotely interchangeable staged reflection gratings. The measured spectral intensity allows to compute the absolute magnitude of the bunch form factor but not its phase which, however, is needed to retrieve the bunch profile. Two phase retrieval methods are investigated in detail: analytic phase computation by means of the Kramers-Kronig dispersion relation, and iterative phase retrieval. Several computational techniques are compared and evaluated in view of their applicability and efficiency. For a large variety of bunch shapes, the time profiles derived from the spectroscopic data are compared with the TDS profiles, and generally excellent agreement is observed down to the 10 fs level. For part of the measurements, two independent CTR spectrometer systems have been available, yielding almost identical bunch shapes. In summary, we demonstrate that using well calibrated and broadband spectroscopy data, a fast and reliable phase reconstruction algorithm leads to bunch profiles competitive to high resolution TDS measurements.

DOI: 10.1103/PhysRevAccelBeams.23.062801

I. INTRODUCTION

The electron bunches in high-gain free-electron lasers (FEL) are longitudinally compressed to achieve peak currents in the kA range which are necessary to drive the FEL process in the undulator magnets. Bunch compression is accomplished by a two-stage process: first an energy chirp (energy-position relationship) is imprinted by off-crest acceleration onto the typically 10 ps long bunches emerging from the electron gun, and then the chirped bunches are passed through magnetic chicanes where the length is reduced to about 100 fs due to the momentum dependent path length difference. A linearization of the energy chirp is achieved by superimposing the accelerating field with a higher harmonic.

Magnetic compression of intense electron bunches is strongly affected by collective effects in the chicanes.

Space charge forces, coherent synchrotron radiation and wakefields have a profound influence on the time profile and internal energy distribution of the compressed bunches. There are considerable uncertainties in the theoretical models and hence measurements of bunch length and shape are of great interest.

Two complementary techniques are applied for determining the time profile of the compressed electron bunches at the soft x-ray FEL FLASH [1]: time-domain measurements with a transversely deflecting radio frequency structure (TDS), and frequency-domain measurements of coherent transition radiation (CTR) with a multichannel spectrometer covering the infrared and THz regime. Electron bunch length determination in the subpicosecond regime by Michelson or Martin-Puplett interferometry of coherent radiation was proposed in 1991 [2] and early successful experiments have been reported in Refs. [3,4]. Since then, this well-established technique has been widely used up to date, see, e.g., [5–8]. Long trains of identical bunches are needed: their average autocorrelation function is sampled, changing the delay between the interferometer arms in small steps. Only a few single shot spectroscopic techniques have been reported so far: using a ten channel polychromator [9], a prism based spectrometer [10] and a
one-stage prototype of our spectrometer [11]. Comparison and benchmarking of the radiation based technique with simultaneous time domain measurements has been so far restricted to the picosecond regime using streak cameras [12].

It is the purpose of this paper to validate CTR spectroscopy as a powerful tool for longitudinal beam diagnostics in a direct comparison to high resolution time domain measurements down to the 10 fs regime. It will be shown that on the basis of broadband (2 orders of magnitude) and precisely calibrated spectroscopic data faithful bunch shape reconstruction using phase retrieval algorithms can be achieved. We present numerous measurements of coherent transition radiation spectra in the infrared and THz regime and compute the length and time profile of the electron bunches from the spectroscopic data by applying the experimental techniques and mathematical methods described in detail in Ref. [13].

The layout of the free-electron laser facility FLASH and the location of the diagnostic devices used is shown in Fig. 1. A detailed description of the facility can be found in Ref. [14] and the literature quoted therein.

II. PRINCIPLE OF BUNCH SHAPE RECONSTRUCTION FROM SPECTROSCOPIC DATA

A. Form factor

The three-dimensional bunch form factor is defined as the Fourier transform of the normalized charge density

$$F_{3D}(k) = \int \rho_{3D}(r) e^{ik \cdot r} d^3 r.$$  \hspace{1cm} (1)

Here $k$ is the wave vector with the relations $k = |k| = 2\pi/\lambda = \omega/c$. The coherent spectral energy density produced by a bunch of $N$ electrons is the product of the spectral energy density produced by a single electron, the number of electrons squared and the absolute square of the form factor:

$$\left[ \frac{d^2 U}{d\omega d\Omega} \right]_{\text{coh}} = N^2 \left[ \frac{d^2 U}{d\omega d\Omega} \right]_{\text{trans}} |F_{3D}(k)|^2.$$  \hspace{1cm} (2)

With the simplifying assumption that the transverse density distribution of the electron bunch is independent of the longitudinal position in the bunch, the 3D form factor is the product of the transverse and the longitudinal form factors:

$$F_{3D}(k) = F_{\text{trans}}(k_x, k_y) F_{\text{long}}(k_z).$$  \hspace{1cm} (3)

The finite electron beam radius of typically 100 $\mu$m (rms) leads to a weak suppression of coherent emission at small wavelengths. This suppression effect, in our case a reduction from 1 at $\lambda \geq 100$ $\mu$m to 0.8 at 10 $\mu$m and 0.5 at 5 $\mu$m [13], is included in the overall response function of the spectrometer system.

Once this correction has been taken care of, we can restrict ourselves to the longitudinal form factor which we write now as a function of angular frequency $\omega = ck \approx ck_z$. The Fourier transformation relations between the normalized longitudinal density distribution $\rho(t)$ (here written as a function of time) and the longitudinal form factor $F$ are

$$F(\omega) = \int_{-\infty}^{\infty} \rho(t) e^{i\omega t} dt.$$  \hspace{1cm} (4)

In general the form factor is a complex-valued function

$$F(\omega) = F(\omega)e^{i\Phi(\omega)}.$$  \hspace{1cm} (4)

If both $F(\omega) = |F(\omega)|$ and $\Phi(\omega)$ were known over the full range of nonvanishing $F(\omega)$, a unique reconstruction of the charge distribution $\rho(t)$ could be achieved by the inverse Fourier transformation:

$$\rho(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)e^{-i\omega t} d\omega.$$  \hspace{1cm} (5)

The spectrometer measures the CTR intensity as a function of frequency, from which the magnitude $F(\omega)$ of the form factor can be derived, using Eq. (2). The phase $\Phi(\omega$) remains unknown.

B. Phase retrieval

To reconstruct the unknown phase $\Phi(\omega)$ from the measured form factor modulus and thus to enable the determination of the bunch profile has been first proposed by Lai and Sievers [15]. Unfortunately, the method has a
fundamental limitation: the unique reconstruction of a function from the magnitude of its Fourier transform is mathematically impossible in the one-dimensional case. The (1-D) phase problem has uncountably many solutions, the precise form of them has been first investigated by Akutowicz [16]. A detailed discussion of the ill-posed nature of the problem can be found in Ref. [17]. Nevertheless, there is at least partial information about the phase hidden in the modulus of the form factor and by using additional constraints on the reconstructed profile, it is possible to find a not only possible but also appropriate solution.

1. Analytic phase retrieval

The well-known Kramers-Kronig dispersion relation between real and imaginary part of the Fourier transform of a real and causal signal can be extended to establish a relation between absolute magnitude and phase, see, e.g., [18]. Lai, Happek, and Sievers [3,19] have adopted this method for the complex form factor of an electron bunch and the reconstruction of asymmetric bunch shapes.

The Kramers-Kronig (KK) phase is computed by means of the following principal-value integral

$$
\Phi_{KK}(\omega) = \frac{2 \omega}{\pi} \int_0^\infty \frac{\ln|F(\omega')| - \ln|F(\omega)|}{\omega^2 - \omega'^2} d\omega'.
$$

A rigorous mathematical derivatization of this formula can be found in Appendix A of Ref. [13]. On experimental data, the integral (6) has to be evaluated by numerical integration methods. The time profile of the bunch is computed by an inverse fast Fourier transformation (IFFT). Since both, the numerical integration of (6) and the IFFT require integration over the full range of nonvanishing $F(\omega)$, most of this range should be covered by experimental data. In any case, extrapolations to low and high frequencies, beyond that range, will be needed.

Because of the above-mentioned ambiguities, the KK phase $\Phi_{KK}(\omega)$ will usually differ from the unknown true phase $\Phi(\omega)$.

2. Iterative phase retrieval

For the iterative phase retrieval we use the Gerchberg-Saxton (GS) algorithm [20] depicted in Fig. 2. Bajlekov et al. [11] have adopted this method for the reconstruction of ultra short bunches from a plasma based accelerator from spectroscopic data.

The $n$th iteration of the loop proceeds as follows. The particle density distribution $\rho_0(t)$ is subjected to a fast Fourier transformation (FFT) to obtain an estimate $G_n(\omega_j) \exp[i\Phi_n(\omega_j)]$ of the complex form factor at the discrete frequencies $\omega_j$. Now a frequency domain constraint is applied: The computed modulus $G_n(\omega_j)$ is retained at those discrete frequencies $\omega_j$ where it agrees with the measured values $F(\omega_j)$ within the estimated experimental error margin $\Delta_j$. If it is outside this margin, $G_n(\omega_j)$ is replaced by $F(\omega_j)$. The modified Fourier amplitudes are thus

$$
G'_n(\omega_j) = \begin{cases} 
G_n(\omega_j) & \text{if } |G_n(\omega_j) - F(\omega_j)| < \Delta_j \\
F(\omega_j) & \text{if } |G_n(\omega_j) - F(\omega_j)| > \Delta_j.
\end{cases}
$$

All phases $\Phi_n(\omega_j)$ are retained.

Next an inverse Fourier transformation is carried out leading to a modified time profile $\rho_n'(t)$. This profile is subjected to a time domain constraint: the particle density is not allowed to assume negative values. If negative values occur, they are replaced by zeroes. This constraint leads to a modified time profile $\rho_{n+1}(t)$ which is then used as input for the next iteration:

$$
\rho_{n+1}(t) = \begin{cases} 
\rho_n'(t) & \text{if } \rho_n'(t) \geq 0 \\
0 & \text{if } \rho_n'(t) < 0.
\end{cases}
$$

The Gerchberg-Saxton loop is terminated when the Pearson correlation coefficient between the profiles $\rho_{n+1}(t)$ and $\rho_n(t)$ has reached a stable level and does not drop any further with increasing number of iterations. The profile $\rho_n'(t)$ is then the reconstructed charge distribution. The number of iterations needed strongly depends on the quality of the starting distribution.

We start the GS loop with a first guess $\rho_0(t)$ of the bunch time profile. Three cases are considered: (a) time profile computed from the measured form factor combined with random phases, (b) analytical model profile and (c) profile reconstructed using the Kramers-Kronig phase. The three alternatives are in detail: (a) Start with a time profile which is computed by combining the measured form factor modulus with a set of initial phases. Following a widely used procedure in crystallography and other fields [21], these phases may be completely random numbers in the
interval $[0, 2\pi]$, varying independently from one discrete frequency to the next. With this rather unphysical choice, the convergence is slow and very many iterations are needed. Additionally, there is a considerable probability to create persistent “echo-peaks” which have to be removed by postprocessing. A considerable improvement, introduced by Fienup [21], is the start with quasirandom phases. Here the computational steps are as follows: In step 1, all phases are put to zero and an IFFT is applied to the measured $F(\omega_j)$. In step 2, a threshold is imposed in the resulting time profile. All points below threshold are put to zero while the points above threshold are replaced by random numbers. In step 3, an FFT is applied to the discrete time distribution generated in step 2. The resulting phases $\phi_1(\omega_j)$ are quasirandom since their origin is a randomized time distribution. Finally, the start profile $\rho_1(t)$ of the Gerchberg-Saxton loop is computed by applying an IFFT to the set of complex numbers $F(\omega_j) \exp[i\phi_1(\omega_j)]$. (b) Start with a suitable model profile. In that case the iterative loop converges rapidly toward a time profile fulfilling all constraints. The optimum start model is found by comparing the measured form factor with the Fourier transform of different analytic model shapes whose parameters are properly adjusted. The “bias” introduced by selecting a specific start model and the impact of different start models on the final result will be discussed in Sec. IV. (c) Start with the profile based on the Kramers-Kronig phase. In contrast to methods (a) and (b), this leads to only one solution.

III. EXPERIMENTAL TOOLS AND METHODS

A. Transversely deflecting rf structure TDS

Using a transversely deflecting radio frequency structure (TDS) to measure the temporal profile of electron bunches is a well established technique proposed and first used [22,23] at SLAC. Owing to its excellent resolution, the TDS is ideally suited for benchmarking CTR spectroscopy as a tool for longitudinal bunch diagnostics. The TDS used at FLASH was built at SLAC [24]. In the description of this time-domain method we concentrate on the properties that are relevant for the comparison of the two techniques. More details on the application and performance of transversely deflecting rf structures at FLASH can be found in [14,25].

The TDS we are using is a disc-loaded travelling wave structure operating at 2.856 GHz in a hybrid HEM mode with transverse electric and magnetic fields. The principle is explained in Fig. 3. The sine curves show the electromagnetic force acting on the electrons. The rf phase is chosen such that the bunch center $\zeta = 0$ coincides with a zero-crossing of the rf wave. This condition holds along the entire axis of the TDS because bunch and phase of the deflecting mode move with almost the same speed.

Each electron is subject to a vertical force which depends on its internal coordinate $\zeta$ along the bunch axis. The resulting slope is transformed into a vertical offset on the view screen due to the $\pi/2$ betatron phase advance between TDS and screen. The vertical axis on the screen is equivalent to a time axis. By projecting the screen image onto this axis one obtains, in the ideal case, the time profile of the bunch. In the vicinity of the zero-crossings the sine curves can be replaced by linear functions and the mapping of the internal bunch coordinate $\zeta$ onto the $t$ axis of the view screen is almost linear.

An essential prerequisite for achieving excellent time resolution is an optimized beam optics in the TDS region. The time resolution of the TDS is derived from measurements without rf field. For our measurements, the rms resolution $\sigma_t$ varied typically between $\sigma_t = 7$ fs and $\sigma_t = 15$ fs. We note that the beam optics required for FEL operation in FLASH degrades the resolution to more than 50 fs.

In principle the TDS is a single-shot device that should be suited to record faithfully the longitudinal charge density profile of a single electron bunch. However, a complication arises since the electron bunches enter the TDS with a nonvanishing correlation between the longitudinal position $\zeta$ and the slice centroid lateral position $\langle y \rangle$ and slope $\langle y' \rangle$. Origin of these correlations may be wakefields, coupler kicks, and spurious dispersion. If we call the vertical coordinate axis at the screen $u$ to distinguish it from the phase space coordinate $y$ of the particles, the correlations result in an additional centroid offset $u_c(\zeta)$ which interferes with the linear $u(\zeta)$ dependence imprinted by the TDS and thus perturbs the streak image on the screen. As a consequence, the image depends on the streak direction in a non-trivial way. The problem can be cured by making measurements with two opposite streak directions [22]. The unperturbed bunch profile and the correlation function can be obtained using a reconstruction algorithm explained in the Appendix. An example is shown in Fig. 4.

The various steps leading from the screen image to the final TDS profile will be outlined briefly. The screen images from the two opposite streaks are cleaned from background and projected onto the vertical $u$ axis. This
yields two raw-data line profiles. To mitigate noise and fluctuations, typically 20 line profiles are averaged for each streak direction. The averaging is impeded by the phase and amplitude jitter of the TDS-RF, leading to a centroid jitter of the streaked profiles and a jitter of the $\zeta \rightarrow u$ relationship. It has to be compensated by equalizing the first and second moments (centroid and rms width) of the single-shot profiles. The two resulting averaged profiles are input to the reconstruction algorithm from which the true longitudinal charge profile $\rho_0(\zeta)$ as well as the intrinsic centroid shift $u_c(\zeta)$ can be recovered. The mean reconstructed TDS profiles obtained in this way are the basis for the comparison with the spectroscopic technique. Under the assumption that $u_c(\zeta)$ does not fluctuate from bunch to bunch, single-shot profiles can be reconstructed from only one streak image by using the mean centroid correlation function. Examples for the fluctuations of the individual bunch profiles are shown in Fig. 5.

B. Spectroscopic method

Transition radiation (TR) is produced on metallized screens, inside the vacuum beam pipe of the FLASH linac, at the 141 m and 202 m positions. The screens are tilted by 45° hence backward TR is emitted perpendicular to the electron beam axis and can be coupled out through diamond windows. The spectral intensity arriving at the spectrometer has to be computed numerically since the setup is in the near-field diffraction regime. A custom made wavefront propagation code is used to determine the fraction of radiation intensity reaching the detector elements. The finite size of the TR screen and diffraction in the optical beam transport system are explicitly taken into account. The procedures are thoroughly discussed in Ref. [13]. One coherent radiation intensity spectrometer (CRISP) is permanently installed inside the linac tunnel at the 202 m position (CRISP-202). The radiation is guided from the TR screen to the spectrometer through a 7.7 m long evacuated optical beam transfer channel (see Figs. 29, 30 of Ref. [13]). For part of the runs, a second spectrometer (CRISP-141) was mounted at the focal plane of an 18.7 m long ultrabroadband THz beamline [26] guiding CTR produced on a screen at the 141 m position of the linac to an outside laboratory. A common feature of both optical beamlines is that they provide—in combination with the ring mirrors in the spectrometer (see Fig. 6)—a point-to-point imaging of the narrow beam spot on the CTR screen onto the pyroelectric detectors. Thereby the diffuse background from synchrotron or wakefield radiation by upstream sources is strongly suppressed.

I. Broadband spectrometer

A detailed description of the multichannel infrared and THz spectrometer CRISP can be found in [27], here we restrict ourselves to a brief overview. The spectrometer is equipped with four consecutive blazed reflection gratings G1 to G4, each spreading the radiation over 30 wavelength channels (see Fig. 6). The grating stages exist in two variants, one for the infrared (IR) regime (5.1–43.5 μm i.e., 6.9–58.8 THz), the other for the THz regime (45.3–432.8 μm i.e., 0.7–6.7 THz).

The spectral response of the pyroelectric detector arrays was measured with tunable infrared radiation at the free-electron laser FELIX [28] in the range from 6 μm to
135 μm. One source of systematic error was the lack of well-calibrated power meters covering the entire wavelength range. A preliminary response function $R_{\text{ref}}(\omega)$, in terms of voltage seen at the ADC channels, of the entire Spectrometer System, comprising CTR source, diamond window, CTR beamline and CRISP spectrometer, has been computed by combining the measured detector response with a wavefront propagation calculation for coherent transition radiation, produced by an infinitely short reference bunch ($F(\omega) = 1$) with a charge $Q_{\text{ref}} = 100$ pC. This preliminary overall response function is shown in Fig. 7 (yellow dots).

The response function is not a smooth function of frequency but exhibits a sawtoothlike structure, caused by the variation of solid angle acceptance within each detector array. The discrete frequencies $f_1, f_2, \ldots, f_{120}$ (in either the IR mode or the THz mode) are the center values of the frequency bins $f_m \pm \delta f_m$ associated with the 120 pyroelectric detectors. The sawtoothlike structure is caused by the variation of solid angle acceptance within each detector array.

The final response function is also shown in Fig. 7 (blue dots). It relates indirectly the spectral transition radiation energy $U_m$ inside the frequency bin $\omega_m \pm \delta \omega_m$ and thus the form factor $F(\omega_m)$ to the ADC input voltage $v_m$ of the corresponding pyroelectric detector [Eq. (9)].

### 2. Determination of form factor

The spectrometer system response to an arbitrary bunch with a different charge $Q$ and a finite length ($F(\omega) < 1$) can be easily written down. The radiation energy $U_m$ in the frequency bin $\omega_m \pm \delta \omega_m$, generated by this bunch, leads to

![FIG. 6. (a) Schematic view of the spectrometer equipped with a bandwidth-limiting grating G0 and four stages of blazed reflection gratings G1 to G4. (b) Layout of a grating stage: reflection grating, ring mirror and line array of 30 pyroelectric detectors. Sketched are the light dispersion and focusing, computed with a ray tracing code, for 5 of the 30 wavelength channels.](image)

![FIG. 7. The response function $R_{\text{ref}}(\omega)$ of the multichannel spectrometer as installed at FLASH, plotted versus $f = \omega / 2\pi$. The discrete frequencies $f_1, f_2, \ldots, f_{120}$ (in either the IR mode or the THz mode) are the center values of the frequency bins $f_m \pm \delta f_m$ associated with the 120 pyroelectric detectors. The sawtoothlike structure is caused by the variation of solid angle acceptance within each detector array.](image)

![FIG. 8. The form factor of a highly compressed bunch, computed after the final calibration. The inset shows the reconstructed bunch shape. A TDS measurement of this bunch has not been possible because the large energy spread exceeded the acceptance of the TDS system.](image)
an ADC input voltage \( v_m(\omega_m) \) and the form factor is then given by

\[
F(\omega_m)^2 = \left( \frac{Q_{\text{ref}}}{Q} \right)^2 \left( \frac{v_m(\omega_m)}{R_{\text{ref}}(\omega_m)} \right) \quad (m = 1, 2 \ldots 120). \tag{9}
\]

Note that this equation must be used twice, in the IR mode and in the THz mode.

The parallel readout of the CRISP spectrometer permits the measurement of the CTR spectrum generated by a single bunch. For the results shown here, the form factors have been derived as averages of 200 single bunches, kicked from consecutive bunch trains at 10 Hz repetition rate. To cover the full spectral range from 0.7 THz to 60 THz, consecutive measurements with the two grating sets are needed. The measurement series with the two grating sets are carried out consecutively and they are individually averaged. The stability of the accelerator is generally quite high and the averaged form factor values are stable within the required data taking time of about 2 minutes. The averaging procedure reduces detector and amplifier noise as well as statistical fluctuations, and it extends the applicability of spectroscopic bunch shape analysis to low bunch charges and large bunch lengths.

Besides detector and electronics noise there exists radiation background which might possibly disturb the form factor measurements. One source is incoherent transition radiation. The intensity ratio of coherent to incoherent transition radiation is

\[
I_{\text{coh}}(\omega)/I_{\text{incoh}}(\omega) = N|F(\omega)|^2
\]

where \( N \) is the number of electrons per bunch. For a typical particle number \( N = 10^9 \) the contribution of incoherent radiation is negligible as long as the magnitude of the form factor stays well above \( 10^{-4} \). Other backgrounds such as wakefield radiation or synchrotron radiation from kicker magnets are negligible since they impinge on the CTR screen as diffuse and widely spread radiation and are thus not imaged onto the small pyroelectric detector elements.

Several data processing steps have to be applied which are illustrated in Fig. 9 where an experimentally determined form factor is shown as function of frequency: (i) Only significant data points are retained with a voltage signal at least 4 standard deviations above noise. (ii) Points with a strong excursion from their adjacent channels are removed. (iii) The data are extrapolated into the frequency range not covered by the spectrometer.

The DC value \( F(\omega = 0) \) is directly related to the total charge of the bunch which is measured independently by toroids and used for the conversion from signal level to form factor according to Eq. (9). The fact that the thus obtained form factors can be smoothly extrapolated to \( F(0) = 1 \) means that the absolute calibration of the entire detection system, comprising the CTR screen, the optical beam transport system and the spectrometer, is very well known. The deviation between extrapolated value and 1 is typically less than about 20% and most probably due to fluctuations of the beam position at the TR screen. To retain \( F(0) = 1 \) together with the smooth extrapolation, the calibration curve is individually re-scaled by a constant factor. The influence of the extrapolations on the reconstructed bunch shapes is discussed in Sec. IV.

\[\text{FIG. 9. An example for a measured form factor modulus } F(\omega) \text{ as function frequency of } f = \omega/2\pi. \text{ The smooth red curve is a linear interpolation of the averaged measured values. The error bars indicate the shot-to-shot fluctuations in the sequence of 200 bunches. The shaded area shows the } > 4\sigma \text{ limit of the noise signal. The extrapolations to low and high frequencies are shown as dotted blue curves. Several data points in the high frequency regime are marked with black crosses. These are severely impeded by detector noise and classified as insignificant.}\]

IV. EXPERIMENTAL RESULTS

Dedicated spectroscopic and time-domain measurements have been made for a large variety of different bunch shapes. The longitudinal charge density profiles obtained by CTR spectroscopy (called “CRISP profiles” for short) will be compared with the “TDS profiles.” The time resolution of the TDS is \( \sigma_t \approx 10 \) fs while CTR spectroscopy is capable of resolving finer structures. To match the resolutions in the comparison of the two methods, the CTR profiles are convoluted with the TDS time resolution function.

A. Benchmarking on form factor level

As a first demonstration of the capacity of our spectroscopic system and to avoid the complex problems related to the phase reconstruction, we compare in Fig. 10 a spectroscopically measured form factor with the Fourier transform of the corresponding TDS charge profile. Since this bunch has a pronounced rectangular shape of duration \( T \), the form factor shows a series of distinct minima with a frequency
spacing of $\Delta f = 1/T$ which are very well resolved by the spectrometer up to about 10 THz. Above this frequency, the TDS Fourier transform shows a steeper drop due to the limited time resolution of our TDS system.

These data demonstrate the outstanding performance of the CRISP spectroscopy system: (i) Broad frequency range (ii) Precise frequency calibration (iii) Fine granularity (iv) Well-calibrated response function $R_{\text{ref}}(\omega_m)$. A comparison of the two measurements in time domain, that is including the phase reconstruction, is shown in Fig. 18 (top left).

**B. Evaluation of different bunch shape reconstruction algorithms**

Keeping in mind that bunch shape reconstruction from spectroscopic data is intrinsically ambiguous (see, e.g., Refs. [13,19]), one has to decide on the aim of the analysis. As known from the TDS measurements as well as from simulations, the bunches in the FEL linac are expected to be compact objects with no or little internal structure. The aim of the phase reconstruction is therefore to find the most compact and strictly positive particle density profile that is compatible with the experimental form factor modulus. In the following we investigate which of the different phase retrieval and pulse shape reconstruction methods yields the desired compact time profiles. Two typical bunch shapes are considered: bunch $B_1$ has an approximately rectangular profile, bunch $B_2$ has a steep rise followed by a smooth drop.

**1. Analytic bunch shape reconstruction**

The Kramers-Kronig (KK) phase is computed by numerical evaluation of the integral (6). The time profile of the bunch is computed by an IFFT, it is shown in Fig. 11.

We note that the KK profiles agree very well with the TDS profiles, except for some unphysical long tails and negative undershoots outside the main peak region.

**2. Iterative reconstruction with quasirandom phases**

For the iterative FFT loop, the measured form factor $F(\omega)$, again including the extrapolations to low and high frequencies, is interpolated and evaluated at the discrete points $f_j$ of a frequency grid with typically 6000 grid points. The sampling in frequency space is dynamically adjusted such that the corresponding time domain grid contains at least 200 sample points over the width of the autocorrelation function. Starting the iterative reconstruction with the measured form factor modulus $F$ and quasirandom initial phases leads to the time profiles depicted in Fig. 12. They exhibit considerable oscillations, which are of course perfectly legitimate and fully compatible with the experimental form factor $F$. Starting with a new seed generally leads to a different time profile. To overcome these unavoidable ambiguities, Pelliccia and Sen [29] propose to repeat the iteration loop many times with different random seeds and average the resulting profiles using an algorithm selecting “similar” shapes. In fact this eliminates the irreproducible structures but on the other hand has a tendency to wash out real features due to the uncertainties concerning a shift and inversion of the time axis. Additionally, the method is extremely time consuming since a large number of slowly converging loops from random start phases have to be calculated. Due to this, and since the method does not offer benefits compared to competing approaches, it is inappropriate for fast online bunch shape monitoring.
3. Iterative reconstruction starting with a model pulse

Various model shapes have been considered for starting the iterative loop, among them rectangle, trapezoid, triangle, Gaussian and step-exponential (i.e., \( g(t) = 0 \) for \( t < 0 \) and \( g(t) = \exp(-\beta t) \) for \( t \geq 0 \)). The Fourier transforms of these models are compared with the measured form factor, and the coefficient of determination \( R^2 \) is utilized to find the optimum start model. For our two sample bunches, the optimum start models are a rectangle and a step-exponential function respectively. The resulting profiles are again in good agreement with the TDS data (Fig. 13). Since the start model is selected by an automation based on the optimum \( R^2 \), it is interesting to check how much the final result is affected by using different and less optimal starts. Rectangular bunches are particularly well suited for studying this impact due to their pronounced structure in the form factor. We demonstrate this for bunch B\(_1\), using the three models depicted in Fig. 14. The Fourier transforms of the non-rectangular models are clearly incompatible with the measured form factor. The effect on the reconstructed shape is obvious, inadequate input models create time profiles with more internal structure. We note that bunches with more smooth form factors are less sensitive to the choice of the start model.

4. Iterative reconstruction using the Kramers-Kronig profile as start

An alternative to starting the loop from analytical model profiles is to use the result based on the Kramers-Kronig phase as input. The motivation to further improve the Kramers-Kronig profile is shown in Fig. 11. The KK profile itself is not always a fully satisfactory representation of the most likely bunch shape since quite often unphysical long tails with negative undershoots are observed. The important constraint that the particle density must be positive cannot be implemented in the analytic method. Replacing the negative values by zero is not a viable option because this may modify the form factor in such a way that it disagrees with measurement. An effective method to enforce observation of the time domain constraints is an ensuing iterative reconstruction starting from the KK profile. Then wiggles and negative undershoots vanish after a few cycles leading to satisfactory results, see Fig. 15.

C. Choice of the optimum algorithm

In summary of the evaluation of different approaches, we conclude that a combination of the analytical Kramers-Kronig phase with a subsequent few cycle iterative
loop to make the particle density strictly positive is the most suitable procedure for a fast and reliable bunch shape reconstruction. Compared to the purely iterative reconstruction based on analytical model profiles it is slightly faster since it avoids fitting various model form factors to the experimental data to select the proper start model. The results of both approaches are identical within the numerical uncertainties of the methods. In contrast to this, iterative reconstruction with random or quasirandom initial phases leads to profiles with random internal structures which cannot be ruled out but violate our demand for the most compact shape compatible with the experimental data. A postprocessing and averaging of a large number of random-start profiles creates additional uncertainties and makes the process very time consuming and inadequate for an online bunch shape monitoring.

D. Form factor at very high and very low frequencies

The CRISP spectrometer permits measurements up to 60 THz and is thus capable of resolving very short time structures, as shown in Fig. 16. In cases like this, where considerably coherent intensity is present at the upper measurement limit, the extrapolation of the form factor to even higher frequencies is a critical issue. One possibility is to assume that the slow drop of the form factor continues beyond 60 THz. This behavior can be parametrized by an inverse power law, \( F(f) = a/f \), with a suitable constant \( a \). The other extreme is a steep drop above 60 THz. Here a parametrization is provided by a Gaussian fall-off \( F(f) = \exp(-bf^2) \). The reconstructed bunch profiles, using the iterative method starting from the KK profile, exhibit both a narrow peak but with significant differences. In case of the \( a/f \) extrapolation the peak current is higher and the leading edge of the bunch is far steeper than in case of the \( \exp(-bf^2) \) extrapolation. The FWHM of the leading spikes is about 10 fs for the Gaussian extrapolation and 4 fs in case of the inverse power law. From the measured data, only an upper limit can be given in such cases.

The extrapolation to low frequencies is generally uncritical. The form factors shown so far can be smoothly extrapolated using a low order Chebyshev polynomial to \( F(0) = 1 \). The situation is different for much longer bunches where \( F \) has already dropped significantly below 1 at the

FIG. 15. Iterative (GS) reconstruction starting from the analytically computed KK time profile.

FIG. 16. The form factor of a highly compressed bunch and two conceivable extrapolations toward higher frequencies: with \( a/f \) or with \( \exp(-bf^2) \) fall-off respectively. The inset shows the reconstructed shapes using the iterative method starting from the KK profile. The bunch is too short to be resolvable by our TDS.

FIG. 17. The form factor of a bunch with a long tail and two conceivable extrapolations toward low frequencies. The blue extrapolation is a smooth continuation of the measured values ignoring the \( F(0) = 1 \) condition, the reconstructed time profile differs significantly from the TDS profile. The constrained extrapolation uses \( F(0) = 1 \), now the reconstructed time profile agrees much better with the TDS profile.
lowest measured frequency bin. One example is presented in Fig. 17. This figure demonstrates the importance of the boundary condition $F(0) = 1$ in the low-frequency extrapolation of the form factor which can only be used as an additional constraint if the absolute response of the spectrometer is calibrated correctly.

E. Comparison of two independent spectroscopic measurement sets

Several measurement series were performed taking data with two CRISP spectrometers simultaneously, at the 141 m and 202 m positions in the linac. Results are shown in Fig. 18. The two independent CTR spectrometers with different input optics yield almost identical bunch shapes, which are moreover in very good agreement with the TDS profiles. These results strongly support the viability of spectroscopic bunch shape diagnostics, using the instrumentation and data evaluation algorithms described in this paper.

V. SUMMARY

We have demonstrated that coherent transition radiation (CTR) spectroscopy is a powerful and highly competitive method for longitudinal beam diagnostics if reliable broadband spectral data are available. Our group has designed and built a 120-channel spectrometer with fast parallel readout, covering the frequency range from 0.7 THz to 60 THz with two remotely interchangeable grating sets which was painstakingly calibrated. The measured spectral intensity allows to compute the absolute magnitude of the bunch form factor but not its phase. Two phase retrieval methods have been investigated in detail: analytic phase computation by means of the Kramers-Kronig dispersion relation, and iterative phase retrieval. Our conclusion is that the best method for determining the longitudinal bunch profile is a combination of both, namely an iterative reconstruction which starts from the analytically computed Kramers-Kronig time profile. This leads to a reliable determination of the most compact charge profile in agreement with the measured spectral intensities. For the first time, simultaneous spectroscopic measurements and high precision time-domain measurements with a transversely deflecting microwave structure (TDS) have been carried out for a large variety of bunch shapes. Excellent agreement has been observed. Part of the measurements were made with two independent spectrometer systems, and also here excellent agreement has been noted with a time resolution of about 10 fs.

Coherent radiation spectroscopy permits real time bunch shape determination. Moreover form factor data and reconstructed profiles might be used for a refined bunch compression feedback. Whereas TDS based time domain monitoring is intrinsically limited by the normal conducting rf resonator to single shot operation at low repetition rate, frequency domain diagnostics using the CRISP spectrometer can be extended to much higher repetition rates. At x-ray FELs with electron energies above 10 GeV, coherent diffraction radiation (CDR) extends to sufficiently high frequencies and can be used for noninvasive high resolution longitudinal profile monitoring with repetition rates up to the MHz regime [30]. This permits simultaneous measurements of all bunches in the long bunch trains that can be accelerated in superconducting linacs.

ACKNOWLEDGMENTS

The authors would like to thank K. Ludwig and P. Smirnov for their technical support, T. K. T. Kövener for his contributions to the spectrometer calibration and E. Hass and P. Winkler for indispensable assistance during the measurement campaigns. The authors acknowledge support from DESY (Hamburg, Germany), a member of the Helmholtz Association HGF.

APPENDIX: ALGORITHM FOR TDS PROFILE RECONSTRUCTION

As mentioned in Sec. III A, an intrinsic transverse correlation along the bunch disturbs the longitudinal profile...
If the original line charge density of the bunch is denoted \( \rho_0(\zeta) \), using two streaks with same shearing amplitudes but opposite sign \( \pm S \) eliminates the interference. With the following procedure, first privately noted down by Loos [31], the longitudinal charge distribution as well as the correlation function can be retrieved. Since our comparison is based on this reconstruction technique, we want to discuss it briefly.

In presence of an intrinsic transverse-longitudinal centroid correlation of the bunch, the mapping function \( \mu \) of the longitudinal bunch coordinate \( \zeta \) to the screen coordinate \( u \) has two contributions: the linear mapping by the TDS field, expressed by the shear parameter \( S \) and the unknown contribution \( u_c(\zeta) \) resulting from the correlation:

\[
\mu : \zeta \mapsto S \cdot \zeta + u_c(\zeta) = u(\zeta)
\]

and similarly for the interval

\[
d\zeta \mapsto S \cdot d\zeta + u_c(\zeta)dy = u'(\zeta)du
\]

A fundamental condition for the reconstruction is that the imprinted streak \( S \) is strong enough that \( \mu \) is an injective function and thus invertible

\[
\mu^{-1} : y \mapsto \zeta_0(u).
\]

If the original line charge density of the bunch is denoted by \( \rho_0(\zeta) \), the measured charge density on the screen is given by

\[
\rho(u) = \frac{dq}{du} = \frac{\rho_0(\zeta_0)}{|u'(\zeta_0)|}.
\]

If two opposite streaks \( \pm S \) (with rf phase difference \( \pi \)) are used they result in two mapping functions \( \mu^\pm(\zeta) \) producing two distinct images \( \rho^\pm(u) \) on the screen.

Without internal correlation, \( S \mapsto -S \) just leads to an inversion \( \rho^+(u) = \rho^-(u) \). With internal correlation, the two streak images are different. The situation is illustrated using a model distribution in Fig. 19(a).

For both streak images, the cumulative charge fractions are defined as follows

\[
q^+(u) = \int_{-\infty}^{u} \rho^+(u_1)du_1 \quad (A5)
\]
\[
q^-(u) = \int_{u}^{\infty} \rho^-(u_1)du_1 \quad (A6)
\]

The inverted direction of integration reflects that the two streaks have an inverted \( \zeta \) to \( u \) mapping, upper resp. lower integration limits refer to the same end of the bunch. Since we requested that the mapping function \( \mu \) is injective, \( q^\pm(u) \) are monotonic and can be inverted leading to the functions \( u^\pm(q) = q^\pm(u)^{-1} \). Since the original ordering of the charge along \( \zeta \) (or time) is not changed by the streaking and subsequent mapping, a specific charge fraction \( q \) refers to the same specific bunch coordinate \( \zeta \). In consequence, one can build the average of \( u^+ \) and \( u^- \)

\[
U(q) = \frac{1}{2} [u^+(q) - u^-(q)] \quad (A7)
\]

in which the contribution of the intrinsic correlations cancels [Fig. 19(b)]. We build again the inverse function \( Q(u) = U(q)^{-1} \), it is the undisturbed cumulative charge fraction along the screen coordinate. Taking the derivative and mapping back to \( \zeta \) with the undisturbed streak \( S \), we get the undisturbed charge density as function of the bunch coordinate as

\[
\rho_0^+(\zeta) = S \cdot Q'(S \cdot \zeta) \quad (A8)
\]
Similarly, we can form the average of $u^+$ and $(+u^-)$ in which the contribution of the two opposite streaks cancel

$$Y_0(q) = \frac{1}{2} [u^+(q) + u^-(q)]$$ (A9)

and get the intrinsic correlation function as

$$u_c(\zeta) = Y_0[Q(S \cdot \zeta)].$$ (A10)

If $u_c(\zeta)$ has been determined, the $\mu^{\pm}(\zeta)$ are known, and $\zeta_0(\mu)$ can be calculated. Using this and Eq. (A4), the undisturbed density distribution can be retrieved from a single stream image as long as $u_c(\zeta)$ does not change.

---

[1] J. Rossbach, J. R. Schneider, and W. Wurth, 10 years of pioneering X-ray science at the Free-Electron Laser at DESY, Phys. Rep. 808, 1 (2019).

[2] W. Barry, An autocorrelation technique for measuring sub-picosecond bunch length using coherent transition radiation, in Proceedings of the Workshop on Advanced Beam Instrumentation, Tsukuba, Japan (KEK-PROC-92-1, 1991), pp. 224-235.

[3] R. Lai, U. Happek, and A. J. Sievers, Measurement of the longitudinal asymmetry of a charged particle bunch from the coherent synchrotron or transition radiation spectrum, Phys. Rev. E 50, R4294 (1994).

[4] H.-C. Lihn, P. Kung, C. Settakorn, H. Wiedemann, and D. Bocek, Measurement of subpicosecond electron pulses, Phys. Rev. E 53, 6413 (1996).

[5] D. Mihalcea, C. L. Bohn, U. Happek, and P. Piot, Longitudinal electron bunch diagnostics using coherent transition radiation, Phys. Rev. Accel. Beams 9, 082801 (2006).

[6] G. Schneider, R. Lai, W. Walecki, and A. J. Sievers, On using the coherent far IR radiation produced by a charged particle bunch to determine its shape—I Measurement with synchrotron and transition radiation, Nucl. Instrum. Methods Phys. Res., Sect. A 396, 283 (1997).

[7] X. Su, L. Yan, Y. Du, Z. Zhang, Z. Zhou, D. Wang, L. Zheng, Q. Tian, W. Huang, and C. Tang, Monitoring of electron bunch length by using Terahertz coherent transition radiation, Nucl. Instrum. Methods Phys. Res., Sect. B 402, 157 (2017).

[8] A. Murokh, J. B. Rosenzweig, M. Hogan, H. Suk, G. Travish, and U. Happek, Bunch length measurement of picosecond electron beams from a photoinjector using coherent transition radiation, Nucl. Instrum. Methods Phys. Res., Sect. A 410, 452 (1998).

[9] T. Watanabe, M. Uesaka, J. Sugahara, T. Ueda, K. Yoshii, Y. Shibata, F. Sakai, S. Kondo, M. Kando, H. Kotaki, and K. Nakajima, Subpicosecond electron single-beam diagnostics by a coherent transition radiation interferometer and a streak camera, Nucl. Instrum. Methods Phys. Res., Sect. A 437, 1 (1999).

[10] T. J. Maxwell, C. Behrens, Y. Ding, A. S. Fisher, J. Frisch, Z. Huang, and H. Loos, Coherent-Radiation Spectroscopy of Few-Femtosecond Electron Bunches Using a Middle-Infrared Prism Spectrometer, Phys. Rev. Lett. 111, 184801 (2013).

[11] S. I. Bajlevkov, M. Heigoldt, A. Popp, J. Wenz, K. Khrennikov, S. Karsch, and S. M. Hooker, Longitudinal electron bunch profile reconstruction by performing phase retrieval on coherent transition radiation spectra, Phys. Rev. Accel. Beams 16, 040701 (2013).

[12] T. Watanabe, J. Sugahara, T. Yoshimatsu, and S. Sasaki, Overall comparison of subpicosecond electron beam diagnostics by the polychromator, the interferometer and the femtosecond streak camera, Nucl. Instrum. Methods Phys. Res., Sect. A 480, 315 (2002).

[13] B. Schmidt, S. Wesch, T. Kövener, C. Behrens, E. Hass, S. Casalbuoni, and P. Schmüser, Longitudinal bunch diagnostics using coherent transition radiation spectroscopy, arXiv:1803.00608.

[14] C. Behrens, N. Gerasimova, C. Gerth, B. Schmidt, E. Schneidmiller, S. Serkez, S. Wesch, and M. Yurkov, Constraints on photon pulse duration from longitudinal electron beam diagnostics at a soft x-ray free-electron laser, Phys. Rev. Accel. Beams 15, 030707 (2012).

[15] R. Lai and A. J. Sievers, Determination of a charged-particle-bunch shape from the coherent far infrared spectrum, Phys. Rev. E 50, R3342 (1994).

[16] E. J. Akutowicz, On the determination of the phase of a Fourier integral. I, Trans. Am. Math. Soc. 83, 179 (1956).

[17] R. Barakat and G. Newsam, Algorithms for reconstruction of partially known, band-limited Fourier-transform pairs from noisy data, J. Opt. Soc. Am. A 2, 2027 (1985).

[18] F. Wooten, Optical Properties of Solids (Academic Press Inc., New York, 1972).

[19] R. Lai and A. J. Sievers, On using the coherent far IR radiation produced by a charged-particle bunch to determine its shape: I Analysis, Nucl. Instrum. Methods Phys. Res., Sect. A 397, 221 (1997).

[20] R. W. Gerchberg and W. O. Saxton, A practical algorithm for the determination of phase from image and diffraction plane pictures, Optik (Stuttgart) 35, 237 (1972).

[21] J. R. Fienup, Phase retrieval algorithms: A comparison, Appl. Opt. 21, 2758 (1982).

[22] P. Emma, J. Frisch, and P. Krejcik, A transverse rf deflecting structure for bunch length and space phase diagnostics, Report No. SLAC LCLS-Technical-Note 00-12, 2000.

[23] R. Akre, L. Benton, and P. Emma, Bunch length measurements using a transverse rf deflecting structure in the SLAC linac, in Proceedings of the 8th European Particle Accelerator Conference (EPS-IGA and CERN, Geneva, 2002), pp. 1882–1884.

[24] O. H. Altenmueller, R. R. Larsen, and G. A. Loew, Investigations of traveling-wave separators for the Stanford two-mile linear accelerator, Rev. Sci. Instrum. 35, 438 (1964).

[25] M. Röhrs, C. Gerth, H. Schlarb, B. Schmidt, and P. Schmüser, Time-resolved electron beam phase space tomography at a soft x-ray free-electron laser, Phys. Rev. Accel. Beams 12, 050704 (2009).

[26] S. Casalbuoni, B. Schmidt, P. Schmüser, V. Arsov, and S. Wesch, Ultrabroadband terahertz source and beamline
based on coherent transition radiation, Phys. Rev. Accel. Beams 12, 030705 (2009).

[27] S. Wesch, B. Schmidt, C. Behrens, H. Delsim-Hashemi, and P. Schmüser, A multi-channel THz and infrared spectrometer for femtosecond electron bunch diagnostics by single-shot spectroscopy of coherent radiation, Nucl. Instrum. Methods Phys. Res., Sect. A 665, 40 (2011).

[28] FELIX laboratory, https://www.ru.nl/felix/ (2020), [Accessed: 30-March-2020].

[29] D. Pelliccia and T. Sen, A two-step method for retrieving the longitudinal profile of an electron bunch from its coherent radiation, Nucl. Instrum. Methods Phys. Res., Sect. A 764, 206 (2014).

[30] N. M. Lockmann, C. Gerth, S. Wesch, and B. Schmidt, Non-invasive THz Spectroscopy for Individual Bunch Current Profile Reconstructions at MHz Repetition Rates, Phys. Rev. Accel. Beams (to be published).

[31] H. Loos (unpublished).