Three-dimensional laser pulse intensity diagnostic for photoinjectors

Heng Li, Ivan V. Bazarov, and Bruce M. Dunham
Cornell Laboratory for Accelerator-based Sciences and Education (CLASSE), Physics Department, Cornell University, Ithaca, New York 14853, USA

Frank W. Wise
Department of Applied Physics, Cornell University, Ithaca, New York 14853, USA
(Received 16 June 2011; published 21 November 2011)

Minimizing the electron-beam emittance of photoinjectors is an important task for maximizing the brightness of the next-generation x-ray facilities, such as free-electron lasers and energy recovery linacs. Optimally shaped laser pulses can significantly reduce emittance. A reliable diagnostic for the laser pulse intensity is required for this purpose. We demonstrate measurement of three-dimensional spatiotemporal intensity profiles, with spatial resolution of 20 μm and temporal resolution of 130 fs. The capability is illustrated by measurements of stacked soliton pulses and pulses from a dissipative-soliton laser.

DOI: 10.1103/PhysRevSTAB.14.112802 PACS numbers: 42.60.Jf, 42.15.Eq, 42.25.Hz, 29.25.Bx

Next-generation x-ray facilities, such as free-electron lasers and energy recovery linacs (ERLs), produce high brightness x-ray beams from diffraction-limited electron beams. The initial electron-beam properties determine the performance of the entire facility, which makes the development of low-emittance electron sources a priority [1]. The beam emittance is a result of the interplay of several phenomena, and depends on a number of factors such as the pulse shape of the photoinjector drive laser [2], the three-dimensional (3D) nature of space-charge forces inside the bunch, the boundary conditions near the photocathode [3], the fields in the radio-frequency (rf) linac cavities, and the aberrations of the electron optics in the gun and downstream. Achieving an ideal 3D electron-beam shape is a matter of active research in the accelerator community: a uniform ellipsoidal beam is the optimal shape when considering linear space-charge forces in free space [4], while a cylindrical shape is known to produce small emittances and is a practical solution pursued in several laboratories [5,6]. However, the optimum intensity profile in a real system generally requires more complicated shapes to achieve the lowest emittance [7]. To experimentally study the effects of the laser shape on beam performance in photoinjectors, a reliable 3D laser pulse intensity diagnostic is required.

Most existing pulse/beam diagnostics measure the field in the space and time domains separately. Second-order autocorrelation is one of the more traditional techniques in the laser field; being simple in its implementation, it, however, can only provide limited temporal and phase information [8]. Frequency-resolved optical gating (FROG) and its successors give both the temporal intensity and phase information, through the spectrogram of the sum frequency generated by the original laser pulse [9]. FROG employs an iterative phase-retrieval algorithm, which works well for most applications. The spectral phase interferometry for direct electric-field reconstruction (SPIDER) technique can also measure the optical field (both the amplitude and phase) by use of a spectral shearing interferometer [10]. Both of these are established techniques for characterization of the full electric field of a light pulse. Typically, a charge-coupled device (CCD) camera is used to display the (time-integrated) transverse intensity profile. While measuring the spatial and time domains separately suffices for many practical cases, the flexibility of the electron generation mechanism in photoinjectors calls for a more general diagnostic capable of providing direct 3D intensity information for the laser pulses incident on the cathode. The phase information is not necessary because the photocathode is not sensitive to the phase. Two-dimensional (2D) spatiotemporal field diagnostics have been developed [11–14] and 3D diagnostics have also achieved limited success [15–17]. Recently, Li et al. developed a simple scheme to measure the 3D laser pulse intensity [18–20], which is essentially a first order noncollinear cross correlation with a CCD camera as a detector. The proof-of-principle experiment has been performed in [18–20], but a number of questions remain regarding the method’s limitations and capabilities. As a first-order autocorrelation technique, it cannot be sensitive to the phase; would a cross correlation between a transform-limited (TL) pulse and the unknown pulse yield more information than an autocorrelation, which washes out all the phase information? What sets the accuracy of the method? Can the method be verified using other

Published by the American Physical Society under the terms of the Creative Commons Attribution 3.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.
techniques and in what cases does it serve as a reliable 3D laser shape diagnostic? What role does data processing play in arriving at the intensity profiles and how can one improve the data analysis? The present work presents a ready-to-use data analysis toolkit and explores the method’s capability in a number of controlled experiments, confirming the accuracy and the practicality of this technique. A quantitative comparison between the measured results and theoretical predictions is presented. The limitations of the method are discussed as well.

We begin with a brief survey of the existing 3D laser pulse diagnostics and follow with the implementation of the 3D intensity measurement. Then we present the measurements of stacked soliton pulses, which offer one practical solution for minimization of electron-beam emittance in photoinjectors [6]. We further illustrate the capability of the 3D diagnostic by measuring the pulse from an all-normal dispersion (ANDi) fiber laser [21], which has a complicated temporal shape. We conclude with a discussion of the limitations of this technique and the outlook for future development.

I. SPATIOTEMPORAL DIAGNOSTICS

Although diagnostics of the temporal pulse shape are well developed, 3D spatiotemporal laser diagnostics have had limited success, in part because of the challenges in their implementation. Unlike the temporal field, which can be mapped to some equivalent one-dimensional representation (such as the spectrum in the frequency domain or intensity in the time domain), or a two-dimensional representation (such as the spectrogram) and easily measured, the three-dimensional field can only be accessed by brute force. Thus, the majority of the methods are based on autocorrelation or cross correlation with an appropriately prepared probe pulse. Di Trapani et al. suggested a cross correlation scheme to image short optical pulses based on noncollinear sum frequency generation; it offers new capabilities to measure the 3D intensity, but the finite bandwidth of the mixing crystal and the noncollinear mixing process limit its spatiotemporal resolution and require extensive data interpretation [15]. Trebino et al. developed a technique based on nonlinear spectral interferometry and, in particular, demonstrated measurement of a pulse at a focus [16]. Trebino and co-workers also demonstrated full three-dimensional intensity and phase information from a single hologram, in which two pulses are cross correlated through a specially made diffractive optical element [17], although the time resolution is limited and the setup is highly specialized. Finally, Li et al. presented a generic method based on noncollinear first order cross correlation [18–20], whose implementation is simple and which is suitable for measurement of a large variety of pulses. As a 3D diagnostic, its accuracy and limitations need to be further investigated. We present this systematic study in a number of controlled experiments.

II. METHOD AND IMPLEMENTATION

This method is essentially a noncollinear first order cross correlation, as illustrated by Fig. 1(a). The pulse under investigation (object pulse) is split into two beams to produce a probe pulse. Ideally, the probe beam should be the transform-limited (TL) version of the original pulse, which was the case in our experiment. These two pulses are combined at a small angle and then detected by a CCD camera. The angle between these two beams is chosen so that the interference fringe spatial period is small but resolvable on the CCD camera. The time delay between the probe and object pulses is adjusted by an optical delay line, and the CCD camera records the time-integrated intensity (fluence) $I(\tilde{r}, \tau)$ as a function of the time delay $\tau$ [20] (we refer to this quantity simply as intensity for the rest of the paper):

$$I(\tilde{r}, \tau) = \int dt |A_o(\tilde{r}, t)\exp[i\phi_o(t) + i\tilde{k}_o \cdot \tilde{r} - i\omega_0 t]|^2 + A_p(\tilde{r}, t - \tau)\exp[i\phi_p(t - \tau) + i\tilde{k}_p \cdot \tilde{r} - i\omega_0 (t - \tau)]$$

$$= I_o(\tilde{r}) + I_p(\tilde{r}) + 2\cos(\omega\tau + \delta(\tilde{r}))$$

$$\times \int dt A_o(t, \tilde{r})A_p^*(t - \delta(\tilde{r}) - \tau, \tilde{r})$$

$$\times \cos(\phi_o(t) - \phi_p^*(t - \delta(\tilde{r}) - \tau)). \ (1)$$

$A(\tilde{r}, t)$ and $\phi(\tilde{r}, t)$ are the slowly varying amplitude and phase of the pulse; the subscripts $o$ and $p$ denote the object and probe pulses, respectively. $I_o(\tilde{r})$ and $I_p(\tilde{r})$ are the intensities of the object beam and the probe beam separately, with the coupling term representing the interference between them. When the probe pulse is short compared to the object pulse, Eq. (1) can be simplified by applying the $\delta$

![Image](https://example.com/image.png)

FIG. 1. (a) Conceptual and (b) schematic implementation of the diagnostic. (a) The initial beam is split to the probe beam and object beam, with adjustable time delays, then they are combined with a small angle on the CCD camera. (b) The experimental setup to measure stacked soliton pulses: ISO, isolator; HWP, half wave plate; PBS, polarization beam splitter; TEL, telescope; BS, beam splitter; BB, beam block; 1D ST, motorized one-dimensional stage; SH1, shutter 1; SH2, shutter 2; CCD, CCD camera.
function approximation, then we obtain the same formula as in [20],

\[ I(\vec{r}, \tau) = I_o(\vec{r}) + I_p(\vec{r}) + 2 \cos \omega [\tau + \delta(\vec{r})] \times \sqrt{\Delta t_p} i_o(\tau, \vec{r}) \sqrt{I_p(\vec{r})}. \] (2)

Here \( C(\tau, r) \) is the amplitude of the coupling term in Eq. (2); it corresponds to the spatial modulation depth of the interference pattern.

The probe and object beams are combined at an angle for practical reasons. If they overlap perfectly with each other, the time delay line will be sensitive to less than a quarter wavelength; considering the moving part of the delay line and mechanical vibrations in the lab, this configuration is overly sensitive and may produce artifacts that would degrade both the temporal and spatial resolutions. With a small angle between the beams, the ambient noise shifts the phase of the fringes but not the envelope [20]. A larger angle corresponds to finer spatial fringes, and finer resolution. In our experiments, the angle is 0.05 rad, which produces fringes with 20 \( \mu \)m period at 1 \( \mu \)m wavelength. These are resolved by the CCD camera, which has pixel size 4.4 \( \times \) 4.4 \( \mu \)m. Thus, our spatial resolution is 20 \( \mu \)m.

The modulation depth \( C(\tau, r) \) should be measured with good fidelity, and this creates several practical considerations for the choice of the CCD camera. Generally speaking, a low-noise and high dynamic-range CCD camera is preferred. The thermal noise from the CCD camera is a known issue; without cooling, the thermal noise from the CCD camera cannot be ignored, and with lower optical intensity increased Poisson noise will appear in the pulse intensity as in Eq. (2). The contrast ratio \( C(\tau, r) \) can be very small for the peripheral part of a pulse; increasing \( A_o(t, \vec{r}) \) increases the contrast ratio, which improves the signal-to-noise ratio achieved by the CCD camera. Practically, the beam spatial intensity distribution is usually a Gaussian, and magnifying the beam by 2 to 3 times will serve this purpose. We magnified the probe beam by 3 times in our experiment.

According to Eq. (2), at each time delay we need to know the intensities of the object beam, the probe beam, and the coupling term. Thus, we record the beam profiles of the object beam \([img_o(\vec{r}, \tau)]\), the probe beam \([img_p(\vec{r}, \tau)]\), and the overlapping beam \([img_o(\vec{r}, \tau)]\) at each time delay. By applying Eq. (3), we can retrieve the object pulse intensity at each time delay.

The envelope of the spatial modulation gives the contrast ratio \( C(\tau, r) \), while the fine spatial modulation depends on the angle between the object and probe beams. In order to extract the contrast ratio envelope from the interference pattern, we need to take the absolute value of the spatial modulation, and then filter out these modulations with a low-pass filter in the spatial frequency domain:

\[ G(\omega) = \frac{1}{\sqrt{1 + (\omega \omega_0)^{2N}}} \] (4)

We use a Butterworth filter [Eq. (4)], which has a flat response in the pass band and thus introduces little distortion to the contrast envelope. In our experiment, we set \( N = 8 \); the bandpass frequency \( \omega_0 \) is determined by minimizing the probe pulse distortion before and after filtering. After filtering, we inverse Fourier transform the filtered spatial frequency back to the space domain to obtain the contrast ratio envelope at a given time delay \( [C^F(\tau, \vec{r})] \). We also filter the intensity of the probe beam to make the data processing consistent. This can be expressed as

\[ C^F(\tau, \vec{r}) = IFFT2[Filter(FFT2[ABS \times [img_i - img_o - img_p(\vec{r}, \tau)]])]. \] (5)

Here the notations ABS, FFT2, and IFFT2 refer to absolute value, two-dimensional fast Fourier transform, and inverse two-dimensional fast Fourier transform. The superscript \( F \) denotes the filtered results.

Whenever a filter is used in data processing, systematic error is introduced. Because of the low-pass nature of the filter, beam profiles with sharp edges can be smoothed out. In our experiment, all the spatial shapes are essentially Gaussian, so that this was not an issue. But for other profiles such as a flattop, or a high-order super-Gaussian, more optimized filtering may present an advantage. The process is repeated for each intensity slice at a given time delay and their combination provides the 3D laser pulse intensity.

We illustrate the method and its capabilities by performing 3D pulse characterization from two different laser systems. The first is a stack of soliton pulses, which is being used at the Cornell ERL photoinjector [2,6].
second is the measurement of the pulse directly from an all-normal-dispersion (ANDi) fiber laser [21], which displays salient temporal features that allows for quantitative comparison with the known pulse shape.

### III. MEASUREMENT OF STACKED SOLITON PULSES

The experimental setup depicted in Fig. 1(b) consists of the soliton and the diagnostics. Soliton pulse evolution happens in the anomalous group velocity dispersion (GVD) regime, where the phases accumulated from the nonlinearity and the dispersion cancel each other exactly; this evolution can also happen when the laser cavity has segments of normal and anomalous GVD and net negative GVD. We constructed a soliton laser from a Yb-doped fiber system. A grating pair inside the cavity provides $\frac{1}{4}$ ps$^2$ GVD, while the fiber dispersion is only around $0.08$ ps$^2$, so that the laser operates in the soliton region. The second-order autocorrelation is shown in Fig. 2, and the spectrum exhibits sidebands that correspond

FIG. 3. Passive beam shaping by two YVO$_4$ birefringent crystals. Their optical axes are oriented at 45° to each other. By varying this angle, more complicated pulses can be produced in the time domain.

FIG. 4. The raw data for the stacked soliton case: (a) the probe beam ($img_p$); (b) the object beam ($img_o$); (c) the overlapping of the probe and object beam when they overlap in the time domain (the interference, $img_1$); (d) zoomed in image of (c); (e) the overlapping of the probe and object beam when they do not overlap in the time domain (the interference, $img_2$); (f) zoomed in image of (e).
to dispersive waves. The full-width at half-maximum pulse duration and spectral bandwidth are around 2 ps and 0.7 nm, respectively; these give a time-bandwidth product of 0.4, which is close to the TL value of 0.32. After producing these pulses, we stacked them together to produce a longer cylinderlike pulse [22].

After the pulses are generated from the soliton laser, the beam is split into the object and probe beams [Fig. 1(b)]. The object beam passes through two birefringent crystals that are configured to produce four replicas of linear vertical and horizontal polarization, as in Fig. 3. Two a-cut YVO₄ crystals with lengths 3.4 and 6.8 mm are configured with optical axes 45° to each other and the laser beam strikes with polarization 45° to the optical axis of the first crystal. The ordinary (o) and extraordinary (e) pulses are separated inside the birefringent crystal because of their group velocity mismatch; a 1-mm YVO₄ crystal causes ~0.8-ps delay between the o pulse and e pulses at 1 μm wavelength. Because the polarization is 45° with respect to the optical axis, the o and e pulse have the same amplitude. After passing through these two crystals, four replica pulses with equal amplitude are produced, and together they stack to a longer pulse, with nearly flattop distribution in the time domain [6,22,23]. As mentioned before, to reduce the error from the retrieval process, the probe beam is magnified by a 1:3 telescope. A one-dimensional motorized stage with 50-nm precision serves as the optical delay line; the time delay is set with 130-fs precision which defines the time resolution. The shutters 1 and 2 in Fig. 1(b) are used to select the probe, object, and overlapping beams separately. The shaped pulse has two orthogonal linear polarizations, and they are detected separately. The half wave plate (HWP) rotates the polarization of the shaped beam, so that only one polarization projects on the probe beam and contributes to the signal on the CCD camera, while the other polarization is orthogonal to the probe beam polarization, and does not produce a signal. A 14-bit CCD camera (Spiricon GRAS20) serves as the detector, which has spatial resolution 4.4 μm. The data acquisition process precedes the data analysis and both are fully automated in MATLAB—collection of 200 data points in the time domain with image resolution 1600 × 1200 (each file is ~1.9 MB) at each slice takes about ten minutes. After applying Eq. (3) to retrieve the object beam intensity, the 3D intensity information is obtained; the data analysis takes around 60 minutes on a quad-core processor with four threads running simultaneously, and it can be further parallelized if more processors are available.

Some sample raw data are shown in Fig. 4. The probe beam is magnified by ~3 times [Fig. 4(a)] compared to the object beam [Fig. 4(b)], so only the center part with adequate intensity overlaps the object beam. When they overlap each other in the time domain, spatial modulations are produced as seen in Figs. 4(c) and 4(d). When the time delay is large, very little spatial modulation is observed [Figs. 4(e) and 4(f)]. The envelope of the spatial

![FIG. 5. The spatial filter: (a) the image before filtering, ABS(img₁ - imgₚ - imgₒ); (b) the image after filtering; (c) the log scale spatial frequency of (a) before filtering; (d) the log scale spatial frequency intensity of (b), after filtering.](image-url)
modulation changes and reflects the pulse intensity profile at each time delay.

Figure 5(a) shows the absolute spatial modulation envelope of Fig. 4(c). Its spatial frequency spectrum is depicted in Fig. 5(c). The sidebands correspond to the modulation fringes, and the dynamic range is $\sim 10^4$, which is limited by the dynamic range of the CCD camera. After applying the low-pass Butterworth filter, the sidebands are suppressed by a factor of $\sim 10^6$ [Fig. 5(d)] and the inverse Fourier transform gives the absolute envelope of the contrast ratio. The object pulse intensity at a given time delay is reconstructed by applying Eq. (3).

By repeating this process at each time delay, slices of the 3D object pulse intensity are collected; combining them, we can get the 3D intensity of the object beam, as shown in the isointensity surfaces plot [Fig. 6(b)], where the false color represents different light intensities. The temporal intensity is obtained by integrating the pulse intensity at each time delay, as shown in Fig. 6(a), where we can see the four nearly equal peaks of the stacked soliton pulses. Each polarization has two peaks and these peaks are separated by $\sim 3$ ps with each other, as expected; their width is $\sim 2.2$ ps which is slightly larger than the 2 ps interferometric autocorrelation measurement, because of the broadening from the cross correlation. The nonzero intensity values between these peaks come from elliptically polarized light, which can come from the misalignment of the birefringent shaping crystals as well as the cutting angle error ($\pm 0.5^\circ$). To verify the result in the spatial domain, we compared the total intensity from the retrieval to the signal on the CCD camera, as shown in Fig. 7.

FIG. 6. Results of measurement of stacked soliton pulses: (a) the temporal intensity distribution obtained by cross correlation; (b) the three-dimensional intensity distribution, plotted as isointensity surfaces.

FIG. 7. The stacked soliton pulses intensity profile: (a) the normalized probe beam intensity; (b) the object beam intensity on the CCD camera; (c) the object beam intensity difference between retrieval process and the intensity on the CCD camera; (d) the probe beam intensity subtract from the object beam intensity; (e) the object beam intensity from retrieval process; (f) the relative error of the retrieval process.
time-integrated intensity of the object beam on the CCD camera [Fig. 7(b)] and the one from the retrieval process [Fig. 7(e)] are normalized; the difference between them is shown in Fig. 7(c), with the relative error shown in Fig. 7(f). For the object part, when the normalized intensity is greater than 0.5, the relative error is smaller than 8% and the root mean square (rms) error is less than 2%. Thus, the diagnostic also gives good results for the stacked soliton pulses in the spatial domain.

IV. MEASUREMENT OF THE PULSE FROM AN ALL-NORMAL-DISPERSION LASER

This diagnostic can also handle pulses with more complicated temporal structure, such as the pulse produced directly from a Yb-doped ANDi fiber laser [21]. Unlike in the soliton regime, where the pulse energy is limited by the area theorem [24], the ANDi regime can support much higher pulse energy. However, the pulses have a large and nearly linear frequency sweep or chirp, and are therefore not TL [21].

The linearly chirped pulses can be dechirped to TL by a grating pair; with $\delta_{0} = 0.33 \text{ ps}^2$ GVD, we dechirped the pulse to $\sim 200 \text{ fs}$, close to the TL [Fig. 8(a)]. These pulses serve as the probe beam, also magnified by a 1:3 telescope to reduce the retrieval error. With a setup similar to the one used for the stacked soliton, we measured its three-dimensional intensity.

The temporal intensity is depicted in Fig. 9(a). The temporal profile looks similar to the spectrum because the pulse is highly linearly chirped. The retrieval result is very close to the spectrum FT including the phase information, which again demonstrates good fidelity of this diagnostics. The small discrepancy comes from the $\delta$ function approximation in Eq. (2). Although the 200-fs probe pulse is indeed much shorter than the $\sim 10 \text{ ps}$ object pulse, and the broadening effect from the cross correlation is negligible ($< 0.01\%$), the phase needs to be considered because only the coherent parts will add together. For this particular object pulse, the linear phase change across 200 fs is around $\pi$. The phase cross correlation between the TL probe and highly linearly chirped object pulses results in a shape similar to the temporal intensity profile of the object pulse. As discussed above, in this situation, the retrieved intensity profile will have the correct qualitative features, but will have quantitative discrepancy from the correct result. This discrepancy is illustrated in Fig. 9(a).

The three-dimensional intensity is depicted as the iso-intensity surfaces in Fig. 9(b). The method successfully recovers the structure present in the pulse. Figure 10 shows good agreement for the spatial domain, even though the probe pulse profile suffered from scratches on the dechirping grating pair, and is far from being a Gaussian. Good results with less than 10% error for the center part of the object beam and $\sim 2\%$ rms [Fig. 10(f)] can still be obtained when the normalized intensity is larger than 0.5.

![Figure 8](image8.png)

**FIG. 8.** The ANDi fiber laser characterization: (a) the second-order autocorrelation, the black curve is the 0-phase spectrum Fourier transform (FT), the red curve is from measurement; (b) the spectrum.

![Figure 9](image9.png)

**FIG. 9.** The ANDi fiber laser results: (a) the temporal intensity distribution; (b) the three-dimensional intensity distribution, plotted in the isointensity surfaces.
V. DISCUSSION AND LIMITATIONS

We measured the 3D intensity of stacked soliton pulses and the pulse produced directly from an ANDi fiber laser based on this method [18–20]. Our results demonstrate good fidelity of this 3D technique in both the temporal and spatial domains, and illustrate the capabilities of the ready-to-use data analysis toolkit.

Although in the results presented here the temporal and spatial parts of the pulse/beam were decoupled, the method can also handle situations that include space-time coupling. Some degradation in spatial resolution is anticipated from the spatial filtering in the case of sharp flattop or super-Gaussian profiles, although we expect the resolution to remain adequate on a scale relevant for photoinjector applications.

There are some limitations to this method. First, it measures the intensity cross correlation. When the probe pulse is much shorter than the object pulse, the broadening effect is negligible, but the phase effect needs to be considered. The cross-correlated intensity has the same features as the object pulse, but needs to be correctly weighted. Second, as previously mentioned, the digital filters introduce some systematic error. The low-pass filter removes certain high spatial frequency, which will affect very sharp spatial edges in the laser beam profile. Despite these limitations, the method is simple and provides reliable 3D laser intensity profiles. We intend to apply this technique in our studies on how the electron-beam emittance depends on the laser pulse intensity distribution. The 3D diagnostic will help minimize the electron-beam emittance by allowing us to control and optimize the laser pulse shape incident on the photocathode.

ACKNOWLEDGMENTS

We wish to acknowledge the support of John Dobbins for help on the automation, Florian Loehl for help on the digital filtering, and Dimitre Ouzounov for the helpful discussion. This work was supported by NSF Award No. DMR-0807731.
