

# Overview of Phase Space Manipulations of Relativistic Electron Beams

Dao Xiang<sup>1</sup>

*SLAC National Accelerator Laboratory, Menlo Park, CA, 94025, USA*

## Abstract.

Phase space manipulation is a process to rearrange beam's distribution in 6-D phase space. In this paper, we give an overview of the techniques for tailoring beam distribution in 2D, 4D, and 6D phase space to meet the requirements of various applications. These techniques become a new focus of accelerator physics R&D and very likely these advanced concepts will open up new opportunities in advanced accelerators and the science enabled by them.

**Keywords:** harmonic generation; emittance and phase space exchange; emittance partitioning; beam conditioning; free-electron laser

**PACS:** 29.27.-a, 29.27.Bd, 29.27.Fh, 41.60.Cr

## INTRODUCTION

The ability to tailor a beam's 6-D phase space is critical to meet the demands of various applications. For instance, in a collider a transverse cavity can be used to imprint a spatial chirp in beam to improve the luminosity by reducing the effective beam sizes at collision [1]; in emittance compensation a solenoid can be used to align the transverse phase ellipses of each beam slices to reduce the projected beam emittance [2, 3]; in free-electron lasers (FELs), the sustained beam-radiation interaction in a long undulator is used to pack the electrons into equally spaced micro-bunches, such that all the electrons radiate in phase that leads to orders of magnitude enhancement in radiation power compared to the spontaneous radiation; in wake field acceleration, the transformer ratio may be significantly enhanced by shaping the electron beam to a distribution that has linearly ramped current [4, 5, 6, 7, 8].

Phase space manipulation is a process to rearrange beam's distribution in 6-D phase space and is of fundamental interest in accelerator physics. In many cases, a simple phase space manipulation may lead to significant enhancement in the performance and capability of accelerator based facilities. For instance, recent studies show that by using a laser to imprint periodic modulation in electron beam phase space, a single-frequency coherent seed can be amplified and further translated to a mode-locked multichromatic output in an FEL [9]. With this configuration the FEL output consists of a train of mode-locked ultrashort pulses that span a wide frequency gap with a series of equally spaced sharp lines, which may potentially allow one to explore the structure and dynamics of a large number of atomic states simultaneously.

In this paper, we give an overview of the techniques developed for tailoring beam distribution in 2D, 4D, and 6D phase space to meet the requirements of various applications. These techniques become a new focus of accelerator physics R&D.

## 2-D BEAM MANIPULATION

In transverse plane, dipoles are used to bend the beam and quadrupoles are used to focus the beam. When they are integrated in a suitable way, they can form a closed loop (an electron storage ring) which allows the beam to circulate for millions of thousands turns. The synchrotron radiation generated when electron beam goes through bending magnets and undulators has proven to be an important tool for physics, chemistry, biology, material science, etc. in the last 50 years. One trend in light source development is to properly design the lattice to reduce the beam emittance to the picometer (pm) level to realize the so-called ultimate storage ring (see, for example [10]). While it

---

<sup>1</sup> dxiang@slac.stanford.edu

appears trivial to control the transverse beam size with quadrupoles, focusing the beam to extremely small size requires special efforts. For instance, a  $\sim 500$  m final focus beam line is needed to focus the electron beam to nanometer (nm) level in linear colliders [11]; another extreme is electron microscope where the aberration is corrected up to the 5th order to provide sub-50 pm resolution which allows one to see the atoms [12].

In longitudinal plane, beam phase space manipulation typically requires a dispersive element. This is because relativistic electrons travel more or less with the speed of light. As a result, they are rigid in a straight section and it is difficult to rearrange them. With a dispersive element, one can force the particles to follow different paths and the beam longitudinal distribution can be readily shaped. In this section, we will focus on several techniques that aim to tailor the beam distribution in longitudinal plane.

## Bunch compression

The process of bunch compression, to the first order, can be described as a linear transformation where the bunch length is reduced while the energy spread (conservation of phase space area) and peak current (conservation of charge) are increased. This is achieved by first accelerating the beam off-crest in rf cavities to establish a correlation between a particle's energy and its longitudinal position, and then sending the beam through a dispersive chicane. The particles with lower energy are bent more in the chicane and they will have a longer path length. On the contrary, particles with higher energy have shorter path length because they are bent less. As a result, the low energy electrons slip back while the high energy electrons catch up, which leads to bunch compression.

The compression factor, to the first order, is approximately  $C = 1/(1 + hR_{56})$ , where  $h$  is the energy chirp and  $R_{56}$  is the momentum compaction of the chicane. It is straightforward to see that there are many combinations of  $h$  and  $R_{56}$  that can provide the same compression factor. The best set of parameters may vary significantly. For instance, in a superconducting machine, typically a small  $h$  and large  $R_{56}$  is used. This is because the chirp will be increased to  $Ch$  after bunch compression and the beam chirp typically needs to be compensated in downstream rf structures. For superconducting machines, running the beam at off-crest phase to compensate for the energy chirp will reduce the energy gain, which will increase the cost of the facility. Therefore, typically a small initial chirp is chosen to reduce the dechirping cost.

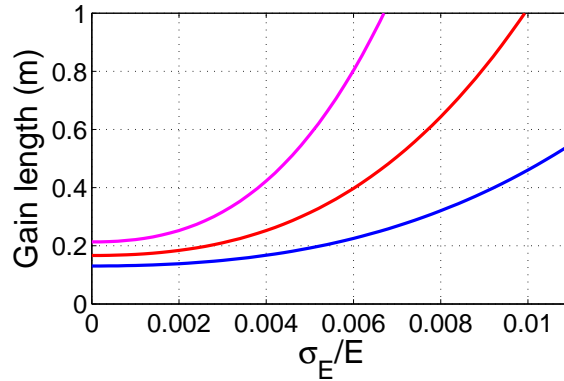
Another concern in bunch compression is collective effects, such as the coherent synchrotron radiation (CSR) effect in the chicane (see, for example [13]). To mitigate CSR effect and loosen the requirement on stability of the rf systems, typically a two-stage bunch compression scheme is used in standard x-ray FELs. Take SLAC's LCLS [14] as an example, for the nominal 250 pC case the beam is compressed by a factor of about 100 in the two bunch compressors to reach  $\sim 3$  kA peak current to drive the x-ray FEL. In the LCLS low charge operation mode (nominally 20 pC) where the collective effects can be greatly mitigated, the beam is normally compressed by a factor of 600 to generate ultra-short x-rays with a few fs duration [15]. Recently, there is also great interest in using beam with even lower charge ( $\sim 1$  pC) to generate attosecond x-rays [16]. However, for this extremely low charge case, the beam will need to be compressed by more than 1000 times, which may put a lot of challenges to the stability of the rf system.

## Bunch decompression

Bunch decompression is a reversal of bunch compression. Though somewhat unusual, bunch decompression may be useful in particular cases where reducing beam energy spread is necessary. For instance, in an energy recovery linac based FEL, after the beam goes through a long undulator to generate intense x-rays, the FEL process tends to increase the beam slice energy spread substantially. While typically this is not a problem at high energy, after the beam is decelerated for energy recovery, the fractional energy spread may become very large and it may result in beam loss if it is beyond the acceptance of the beam line. One solution to reduce the fractional energy spread is to reduce the beam slice energy spread through bunch decompression before it is decelerated [17].

Bunch decompression may be beneficial to FELs driven by laser plasma accelerators (LPAs). The beam from a LPA typically has a high peak current and relatively large energy spread on the  $\sim 1\%$  level. Also the bunch is so short that the slippage length may be comparable to the bunch length. Take the FEL planned at LBNL as an example [18], the representative beam parameters at a moderate energy (500 MeV) are as follows:  $\epsilon_{nx,ny} = 1$   $\mu\text{m}$ ,  $Q = 200$  pC, FWHM bunch length of about 10 fs, peak current of 20 kA and relative energy spread of 1%. Given the FEL wavelength at 31 nm, the slippage length in the undulator that has 220 periods is about 7  $\mu\text{m}$ , which is actually larger than the bunch

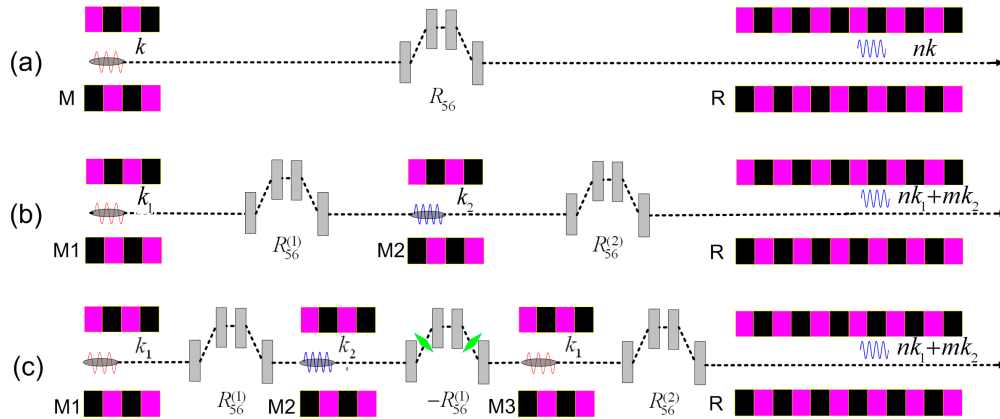
duration. This means that after some distance in the undulator, there will be no electrons to support the FEL growth. The FEL gain lengths for various beam peak currents and energy spreads calculated using Xie's formulae [19] are shown in Fig. 1. With the nominal parameters where beam peak current is 20 kA and energy spread is 1%, the gain length is about 0.5 m. If a chicane is used to decompress the beam by a factor of 2, the beam peak current will be reduced to 10 kA, and accordingly the energy spread is reduced to 0.5%. In this case the gain length is reduced to 0.3 m. Further reducing the peak current to 5 kA may reduce the gain length to below 0.3 m. So bunch decompression may be used to reduce beam energy spread to reduce the FEL gain length and increase the bunch duration to support FEL interaction in longer undulators (see, for example [20]).



**FIGURE 1.** FEL gain length for various energy spreads and peak currents: 20 kA (blue); 10 kA (red); 5 kA (magenta).

### Bunch compression with laser induced energy chirp

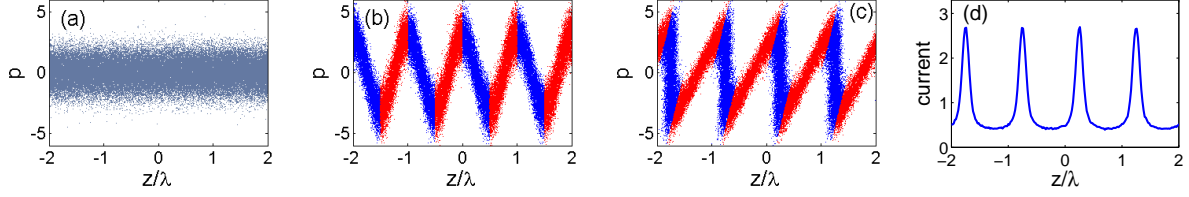
Conventional bunch compression with rf cavities and chicanes allows one to control the beam longitudinal phase space at the micron level. Replacing the rf cavity with lasers will allow manipulation of the beam phase space at nm level. One important application of bunch compression with laser induced chirp is seeding FELs to generate fully coherent radiation. The classic laser seeding technique is called high-gain harmonic generation (HGHG), which is a single modulator-chicane system (see Fig. 2a).



**FIGURE 2.** Various laser seeding schemes for generating fully coherent radiation in FELs: (a) HGHG; (b) EEHG; (c) TMC. "M" stands for modulator and "R" stands for radiator.

In the HGHG scheme [21, 22], a laser with wave number  $k$  is first used to interact with the beam in a short undulator, called modulator, to generate sinusoidal energy modulation in electron beam (Fig. 3b). After passing through a small chicane, half of the particles that have negative energy chirp (blue particles in Fig. 3b and Fig. 3c) will be compressed, while the other half that have positive energy chirp (red particles in Fig. 3b and Fig. 3c) will be decompressed. As a result of this transformation, energy modulation is converted into density modulation and the beam density

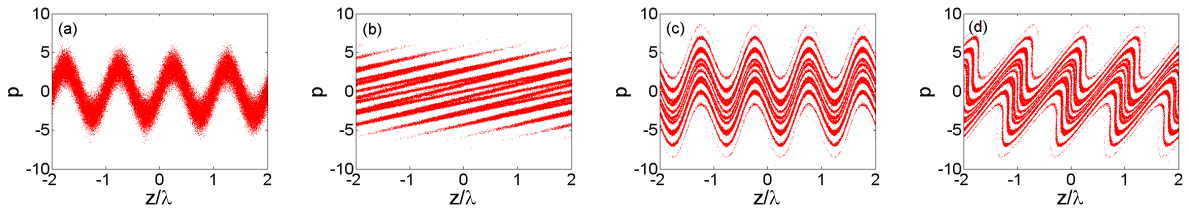
now consists of many bumps equally separated by the laser wavelength (Fig. 3d). Fourier transform of this current distribution yields considerable higher harmonic contents. The density-modulated beam is further sent through a long undulator, called radiator, which is tuned to the harmonic frequency of the seed laser to further amplify the harmonic radiation.



**FIGURE 3.** Basic physics of HGHG. Beam longitudinal phase space before laser modulation (a); after laser modulation (b); after the chicane (c); final current distribution (d). The horizontal axis is the beam longitudinal position normalized to the laser wavelength and the vertical axis is particle's energy deviation with respect to the reference particle normalized to the rms slice energy spread of the beam.

The main limitation of this classic technique is that the frequency up-conversion efficiency is relatively low. Analysis shows that the width of the current bump after the chicane normalized to the laser wavelength approximately equals the ratio of energy spread to energy modulation. Therefore, generation of the  $n$ -th harmonic typically requires the energy modulation to be  $n$  times the energy spread, so high harmonic number requires large modulation. But a large energy modulation is associated with a substantial growth in beam energy spread and the beam quality is significantly degraded. So the harmonic number is typically limited to below 10 in HGHG. As a result, multiple stages of HGHG FELs are needed [23] in order to generate coherent x-rays from a UV laser.

To generate high harmonics with small energy modulation, recently a technique called echo-enabled harmonic generation (EEHG) was proposed, developed and first demonstrated at SLAC [24, 25, 26, 27]. EEHG is a double modulator-chicane scheme (see Fig. 2b), in which two lasers with wave numbers  $k_1$  and  $k_2$  are used to generate high harmonics at the wave number  $nk_1 + mk_2$ , where  $n$  and  $m$  are integers. In the EEHG scheme, the beam is first energy modulated by a laser with wave number  $k_1$  in the first modulator (Fig. 4a) and then sent through a chicane with strong momentum compaction after which the modulation obtained in the first modulator is macroscopically smeared; simultaneously, complicated fine structures are introduced into the phase space of the beam (Fig. 4b). A second laser with wave number  $k_2$  ( $k_2$  can equal  $k_1$ ) is used to further modulate the beam energy in the second modulator (Fig. 4c). After passing through a second chicane the echo signal then occurs as a recoherence effect, like an echo (Fig. 4d). The key advantage of echo technique is that very high harmonics (50 ~ 100) can be generated with relatively small energy modulation, which allows the generation of soft x-rays directly from a UV seed laser.

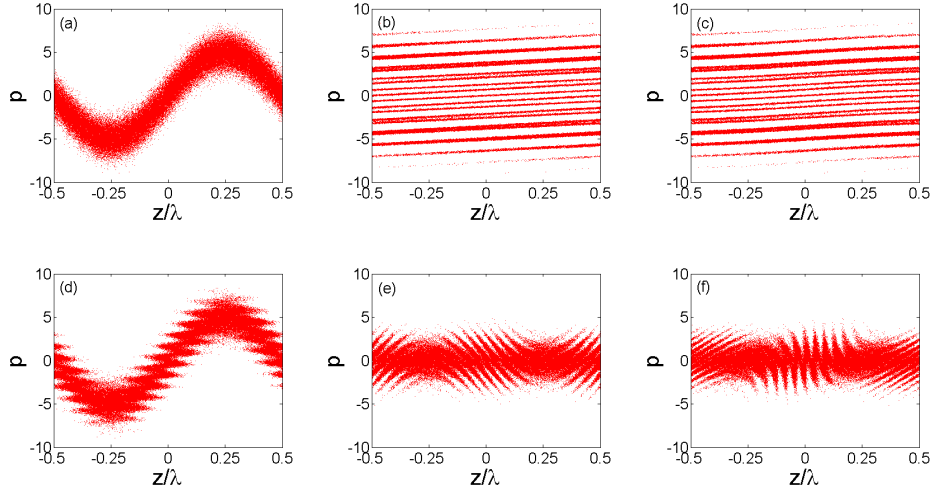


**FIGURE 4.** Evolution of the longitudinal phase space in EEHG scheme. (a) after the first modulator; (b) after the first chicane; (c) after the second modulator; (d) after the second chicane. In this example, the two lasers have the same wavelengths.

To push the harmonic number to beyond 100, very recently a novel seeding scheme that uses a triple modulator-chicane (TMC) system has also been proposed [28]. This scheme is schematically shown in Fig. 2c where the first and second chicanes have opposite momentum compaction and the first and third lasers have the same wavelength and  $\pi$  phase shift. To illustrate the physics behind the TMC scheme, the evolution of the beam longitudinal phase space is shown in Fig. 5. For simplicity, in Fig. 5 the lasers have the same wavelength in the three modulators and only the phase space within one wavelength region is shown.

The beam phase space after interaction with the first laser is shown in Fig. 5a where the energy modulation is 3 times larger than the beam slice energy spread. Similar to the EEHG scheme, separated energy beamlets are generated (Fig. 5b) after the beam passes through the first chicane with a moderate  $R_{56}$  as a result of bunch decompression with a decompression factor  $D$ . After interaction with the second laser of which the energy modulation is 20 times smaller

than beam slice energy spread, the beam phase space evolves to that in Fig. 5c. Because the energy modulation is much smaller than beam's initial energy spread, it is actually very difficult to see the difference between Fig. 5b and Fig. 5c. The second chicane with opposite momentum compaction ( $R_{56}^{(2)} = -R_{56}^{(1)}$ ) compresses the beamlets by a factor of  $D$  and amplifies the modulation imprinted in M2. The resulting beam phase space is shown in Fig. 5d which is similar to that in Fig. 5a. If no modulation is imposed in M2, the second chicane will restore beam phase space to the same distribution as that before the first chicane, because the transfer matrix from the exit of M1 to the exit of the second chicane is unity. With a small energy modulation in M2, the second chicane will transform the beam phase space to a distribution similar to that before the first chicane with the presence of energy modulation from M2 superimposed on the modulation from M1 (Fig. 5d).



**FIGURE 5.** Evolution of the longitudinal phase space in TMC scheme: (a)-After the first modulator; (b)-After the first chicane; (c)-After the second modulator; (d)-After the second chicane; (e)-After the third modulator; (f)-After the third chicane.

The laser in M3 is chosen to give the beam the same modulation amplitude as that in M1, but with  $\pi$  phase shift, so that the overall energy modulation in M1 is canceled in M3. After the cancelation, the modulation from M2 becomes dominant, as shown in Fig. 5e. The wavelength of the modulation in Fig. 5e is roughly  $D$  times shorter than that in M2. A third chicane with small  $R_{56}$  further converts the energy modulation into density modulation (Fig. 5f). By integrating the modulators and chicane in a clever way, the 22nd harmonic is generated with a modulation in M2 that is 20 times smaller than beam slice energy spread and the final energy spread growth is only about 15%.

The unique advantages of TMC scheme that only a small energy modulation is needed in M2 and the second chicane compresses the modulation imprinted in M2 to shorter wavelength opens new opportunities for using low power high-order harmonic generation (HHG) source (see, for example [29]) at short wavelength to seed x-ray FELs. As an example, in [28] the feasibility of generating significant bunching at 1 nm and below from a low power (100 kW) HHG seed at 20 nm assisted by two moderate power (100 MW) UV lasers at 200 nm while keeping the energy spread growth within 40% has been shown. The supreme up-frequency conversion efficiency of the TMC scheme together with its unique advantage in maintaining beam energy spread opens new opportunities for generating fully coherent x-rays at sub-nanometer wavelength from external seeds.

In addition to seeding FELs, bunch compression with laser induced chirp also serves as a standard method to generate optical micro-bunches for net acceleration in direct laser accelerators and dielectric laser accelerators. In the conventional rf accelerators, the electron bunch is much shorter than the wavelength of the rf field, and therefore a monoenergetic beam can be routinely obtained. However, obtaining a monoenergetic beam in laser accelerators is not trivial, because typically the electron bunch length is much longer than laser wavelength and different particles would see different phases (some get accelerated while others are decelerated). One promising way to obtain monoenergetic beams in laser accelerators is to first use a laser to generate optical microbunches, and then put the bumps at the acceleration phase so that most of the particles see more or less the same field which will lead to net acceleration [30].

## RF and optical linearizer

It is well-known that linearizing the beam longitudinal phase space with a harmonic rf cavity can improve the performance of a bunch compressor [31, 32]. Let's assume a beam with energy  $E_i$  is accelerated in two rf linac structures where the main acceleration structure has a wave number at  $k_s$  and the harmonic structure used to linearize the longitudinal phase space has a wave number at  $k_x$ . The peak energy gain of each structure is assumed to be  $E_s$  and  $E_x$ , and the phase relative to the accelerating peak of the waveform is assumed to be  $\phi_s$  and  $\phi_x$  ( $0 < \phi_s < \pi/2$  for obtaining a negative chirp for bunch compression with a simple 4-dipole chicane). The energy of a particle at longitudinal position  $z$  (bunch head at  $z > 0$ ) with respect to the reference particle can be written as,

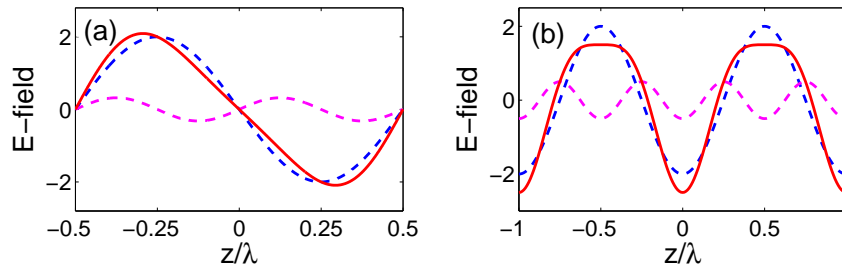
$$E(z) = E_i + E_s \cos(\phi_s + k_s z) + E_x \cos(\phi_x + k_x z) \quad (1)$$

Typically the phase of the harmonic cavity is set at the decelerating phase ( $\phi_x \approx \pm\pi$ ) to compensate for the nonlinear energy chirp, and under this condition the relative energy deviation can be written as,

$$\frac{\Delta E(z)}{E_0} = -\frac{E_s k_s \sin \phi_s}{E_0} z + \frac{1}{2} \frac{E_x k_x^2 - E_s k_s^2 \cos \phi_s}{E_0} z^2 + \Theta(z^3) + \dots \quad (2)$$

where  $E_0 = E_i + E_s \cos \phi_s - E_x$  is the energy of the reference particle. From Eq. (2) one can see that by properly choosing the voltage of the harmonic cavity, the energy chirp of the beam can be linearized up to the 2nd order. Roughly speaking, the required voltage for the harmonic cavity scales as  $1/n^2$  where  $n$  is the harmonic ratio. For instance, when an X-band structure is used to linearize the curvature developed in an S-band linac structure, its voltage is roughly 1/16 of the S-band structure. When the second order transportation effect in a chicane is taken into account, the required linearizing voltage is slightly different from the simplified discussion above, but the physics is still similar.

It is worth mentioning that in addition to compensating for the nonlinear curvature, the harmonic cavity has also been used to reduce the slope of the rf cavity voltage in storage rings to lengthen the equilibrium bunch length to increase the beam life time (see, for example [33]). Similarly, these techniques may be applied at optical wavelengths to enhance the performance of seeded FELs [34, 35] and inverse FELs [36, 37] (IFEL) where a high power laser is used to accelerate the electron beam in an undulator. Fig. 6a shows how a laser harmonic (with frequency at  $2\omega$ ) can be used to linearize the sinusoidal modulation from a laser at the fundamental frequency.



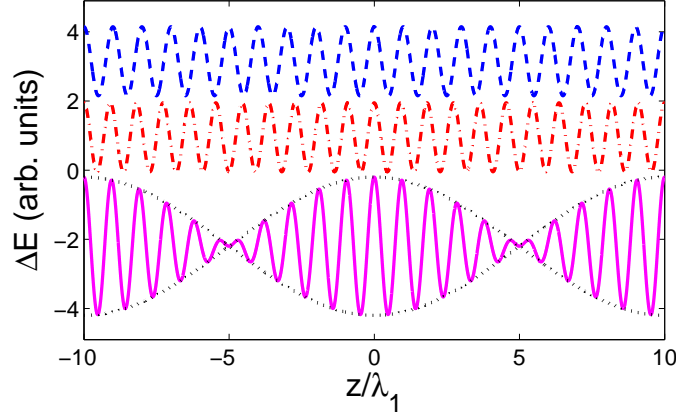
**FIGURE 6.** Energy modulation with a laser at  $\omega$  (blue line), at  $2\omega$  (magenta line), and combination of the two energy modulations (red line).

In Fig. 6a it can be clearly seen that the modulation around the zero-crossing region is much more linear than that from a laser individually. Fig. 6b shows that by shifting the phase of the laser harmonic field by  $\pi/2$ , it can also be used to flatten the  $E$ -field. Combining these two techniques lead to an IFEL with enhanced performance where most of the particles are trapped and accelerated with the same gradient. In this case one can obtain a high quality beam with small energy spread in a laser accelerator.

## Narrow-band THz emission from laser modulated beam

In addition to up-converting the frequency of the lasers to generate fully coherent x-rays, one may also down-convert the frequency of the lasers to generate THz radiation. This can be achieved through manipulation of the beam phase space as well. For example, recently a scheme that consists of two undulators and one chicane has been proposed to generate narrow-band THz radiation [38]. In the first undulator a laser is used to generate energy modulation in the

beam phase space. After interacting with the second laser, the beam phase space consists of fast modulation at the sum frequency and a slow modulation at the difference frequency (see Fig. 7). After passing through a chicane, the energy modulation at the difference frequency can be converted to density modulation. If the frequencies of the two lasers are close to each other, the difference frequency will be much smaller than the laser frequency. In this case the relativistic electron beam is used as the nonlinear medium to down-convert the frequency of two optical lasers to THz range, and therefore one can generate long-scale periodic structures in electron beam through short-scale laser modulations.



**FIGURE 7.** Energy modulation with a laser at  $\lambda_1$  (blue line) and at  $0.9\lambda_1$  (red line); a combination of the two energy modulations (magenta line); the black dotted line illustrates the envelope of the slow modulation at the difference frequency of the two lasers.

Very recently, this technique has been demonstrated at SLAC's NLCTA where density modulation around 10 THz was generated by down-converting the frequencies of an 800 nm laser and a 1550 nm laser [39]. One of the many advantages of this technique is the flexibility it offers to tune the central frequency of the modulation, which can be achieved through tuning of laser wavelengths, beam energy chirp, and chicane momentum compaction. In principle, this allows one to generate coherent narrow-band THz radiation covering the whole THz range.

## 4-D MANIPULATION

For 4-D beam manipulation, one or more elements that couple the beam dynamics in two different planes are needed. For example, solenoids and skew quadrupoles couple the beam dynamics in  $x$  and  $y$  planes; doglegs and rf transverse cavities couple the beam dynamics in  $x(y)$  and  $z$  planes. In this section, we will discuss how a combination of these coupling elements allow one to manipulate the beam distribution in 4-D phase space.

### RF Transverse cavity

A TCAV (also called deflecting cavity or crab cavity) is an rf structure operating in the transverse mode (TM01, TM10, or TM11). When a beam passes through a TCAV at the zero-crossing phase, the TCAV (assuming TM01 mode) imprints a vertical angular kick on the beam that varies linearly with the longitudinal position ( $y' \propto t$ ). After about 90 degrees of phase advance in the vertical plane, the angular distribution is converted to a spatial distribution, and the vertical axis on some screen downstream of the TCAV becomes the time axis ( $y \propto t$ ). The absolute time measurement is calibrated by scanning the TCAV phase and recording the vertical beam centroid motion on the screen. Its high-temporal-resolution and simple calibration make a TCAV well suited for the absolute measurement of beam longitudinal density distributions (see, for example [40, 41, 42]).

It is straightforward to show that in bunch length measurement with a TCAV the vertical beam size on the screen downstream of the TCAV is,

$$\sigma_y = \sqrt{\sigma_{y0}^2 + r_{12}^2 k^2 \sigma_{z0}^2}, \quad (3)$$

where  $\sigma_{y0}$  is the beam size on the screen when the TCAV is off,  $k = 2\pi eV/\lambda_{rf}E$  is the dimensionless kick strength of the TCAV and  $r_{12}$  is the angular-to-spatial element of the transfer matrix from the TCAV to the screen. In order



to unambiguously determine the temporal profile of the beam, one may require the beam size on the screen to be dominated by the kick imprinted in the TCAV. Therefore, the resolution ( $\Delta t = \sigma_y'/k$ ) is limited by the intrinsic beam emittance and the available dimensionless kick of the TCAV. Currently, TCAVs have achieved a resolution of 10~20 fs in measuring ultrashort bunches [40, 41].

From Eq. (3) it is easy to realize that a stronger kick from a TCAV provides a higher temporal resolution. This can be achieved by increasing the voltage and frequency of the TCAV. For instance, given the same voltage, an X-band TCAV provides 4 times higher resolution than an S-band TCAV. An X-band TCAV with maximal voltage of about 50 MV will be installed in FY13 at LCLS to push the resolution to 1 ~ 2 fs [43]. It is also possible to dramatically increase the frequency of the streaking field, say  $\times 1000$ , for which case a laser operating in TEM01 mode is used as the optical deflector. Note, the fact that the kick is only constant around the zero-crossing limited the application of an optical deflector to the case where the bunch is much shorter than the laser wavelength. To improve the dynamic range of the measurement, an 'optical oscilloscope' has been recently proposed where in addition to an optical laser, a TCAV is used to streak the beam in an orthogonal direction [44]. The outcome of such a configuration is a 2-D map which shows the overall bunch shape as well as the fine structures. In principle, with a high power laser the resolution of this method can be pushed to sub-femtosecond.

A more exotic way to improve the temporal resolution of a TCAV is to remove the  $\sigma_{y0}$  term in Eq. (3), namely, making the transverse beam distribution on the screen independent of the beam transverse emittance. In [45] it is shown that by introducing a special chicane with non-zero  $R_{54}$  upstream of the TCAV, particle's initial longitudinal position can be exactly mapped to the final transverse position, which overcomes the fundamental resolution limit arising from the beam intrinsic emittance. The resolution of this longitudinal-to-energy mapping technique is only limited by second order effects and sub-fs resolution may be readily achieved.

## Transverse-to-longitudinal emittance exchange

Another important application of TCAV is for transverse-to-longitudinal emittance exchange (EEX). For a beam line that couples the beam dynamics in  $x$  and  $z$  planes, analysis shows if its 4 by 4 transfer matrix has such a form that it is 2 by 2-block antidiagonal, then a particle's final transverse coordinates will only depend on its initial longitudinal coordinates, and vice versa. As a result, beam's transverse and longitudinal degrees of freedom will be exchanged after the beam passes through the beam line. The first beam line for transverse-to-longitudinal EEX proposed in 2002 consists of a TCAV in the center of a chicane [46]. This scheme is easy to implement, but the exchange is not complete. Later a complete EEX beam line was found in which the TCAV is put between two identical doglegs [47]. However, this scheme introduces offset in beam trajectory which might be undesirable in some cases. Recently a chicane type exact EEX beam line was also proposed [48] where a pair of quadrupoles is used to form a negative unity transfer matrix for the transverse plane in the chicane. The negative unity section reverses the dispersion of the first half of the chicane, which is optically equivalent to flipping the sign of a dogleg. This beam line also provides exact EEX and may be easier to implement because it doesn't introduce offset, and turning off the quadrupoles and TCAV makes the beam line a simple chicane.

One interesting feature of the transverse-to-longitudinal exchanger is that the initial beam timing jitter and energy jitter will be converted into transverse position and angular jitter, and vice versa. In general, a timing jitter on the order of bunch length and an energy jitter on the order of the beam slice energy spread will result in transverse position and angular jitter on the order of the transverse beam size and divergence. While the absorption of timing and energy jitter in EEX may be desired for some facilities, the resulted transverse jitter may be undesired for the facilities that require good pointing stability.

Another interesting scenario in transverse-to-longitudinal EEX is to replace the rf TCAV with a TEM10 mode laser, the so-called laser assisted EEX [49]. The advantage of using a laser to perform EEX is that instead of exchanging the emittance for the whole bunch, one can exchange the emittance for only part of the electrons. This is because the TEM10 laser gives beam a sinusoidal kick while the EEX requires a linear kick. So only the electrons around the zero-crossing will have effective EEX. For these particles, the rms length is only a fraction of the laser wavelength. Therefore it has an extremely small longitudinal emittance. After EEX, it will have an extremely small transverse emittance. With an extremely small emittance beam, one may operate an x-ray FEL at lower energy, which reduces the size and cost of the facility.

One important application enabled by transverse-to-longitudinal EEX is chirp-free bunch compression [50]. As discussed in Sec. II, typically an energy chirp is required for bunch compression, which puts certain constraints on the



linearity of the beam longitudinal phase space. In the chirp-free bunch compression scheme, an EEX beamline is first used to convert  $z$  to  $x$ ; then a telescope beam line is used to demagnify the beam in  $x$ ; finally a second EEX beam line converts  $x$  back to  $z$  and the beam is compressed.

Another application of transverse-to-longitudinal EEX is to tailor beam's longitudinal distribution by shaping beam's initial transverse phase space. For example, at FNAL, a mask is used to generate transversely modulated beamlets, and then with EEX technique, these transverse beamlets are converted into longitudinal sub-ps bunch train [51]. With a special mask, one may also generate a beam with linearly ramped current [52]. It is also straightforward to convert these transverse beamlets into separated energy bands which may be used for generation of high harmonics in FELs [53].

It is worth mentioning that in a typical EEX setup, the final longitudinal coordinates  $(z, \delta)$  depend on both the initial transverse position and transverse divergence. While it is easy to shape the  $x$  distribution using masks, shaping  $x'$  distribution could be technically challenging. In practice one needs to properly design the lattice to make beam's final longitudinal coordinates dominated by initial  $x$  rather than  $x'$ . However, the cross-term dependence may set a limit on how well the phase space can be tailored. For advanced surgery to the phase space, it is desirable to eliminate the dependence on  $x'$ . Further studies show that this can be achieved by adding quadrupoles before and after an EEX beamline [48]. By suitable lattice matching, one can achieve phase space exchange (PSEX) in which both the beam emittance and coordinates are exchanged, i.e. mapping  $x$  to  $z$ ,  $x'$  to  $\delta$ ,  $z$  to  $x$  and  $\delta$  to  $x'$ . Exact PSEX may have wide applications. For instance, by mapping  $z$  exactly to  $x$ , beam's temporal profile can be easily measured with a view screen. By mapping  $x$  exactly to  $z$ , to the first order, one may generate electron beam with arbitrary current using proper transverse masks.

### **Transverse-to-transverse emittance exchange**

EEX in transverse planes is also possible, and is conceptually simpler than the transverse-to-longitudinal EEX. For instance, a solenoid with suitable strength that rotates the beam by  $n\pi/2$  ( $n$  is an odd number) naturally exchanges the beam emittance and coordinates in  $x$  and  $y$  planes. In fact, a beamline entitled 'beam rotator' [54] which consists of 5 skew quadrupoles was proposed in 1972 to effect the transformation  $x \rightarrow y$  and  $y \rightarrow x$ . Such a beamline also effectively exchanges the beam emittance in  $x$  and  $y$  planes.

$X$  to  $y$  EEX was envisioned in the 1980's for LEP injection where the injection oscillation may be reduced by transferring the smaller vertical emittance into horizontal plane [55].  $X$  to  $y$  EEX has also been proposed to convert a circular accelerator into 'Möbius operation' for producing round beams with equal emittances in  $x$  and  $y$  [56]. Specifically, a beamline is used to exchange the horizontal and vertical motions on each passage such that the horizontal betatron motion on one turn becomes vertical on the next turn and vice versa. In this case, radiation damping affects  $x$  and  $y$  in a similar way and one may obtain a beam with full emittance coupling.

### **flat-to-round and round-to-flat transformation**

It should be pointed out that EEX can only exchange the emittance among the eigen-planes, but it does not change them. In some applications, it is desirable to change them. For instance, a round beam is preferred in FELs, because the fundamental mode has the highest growth rate and a round beam is a perfect match to the TEM00 mode radiation. But for linear colliders, a flat beam with large aspect ratio is needed to mitigate the beam-beam effect in the IP. So there is a great need to convert a round beam with equal emittance in  $x$  and  $y$  into a flat beam with large emittance ratio, and vice versa.

The flat-to-round and round-to-flat transformation was first introduced by Derbenev [57] and further developed in many publications [58, 59]. The original motivation for flat-to-round transformation is to enhance the efficiency of high energy electron cooling. In this application, a flat beam is converted into an angular momentum dominated round beam with a triplet of skew quadrupoles; and the angular momentum is eliminated with the entrance field of a solenoid and the cooling occurs inside the solenoid. Later it was realized the apparent emittance of the beam inside the solenoid can be made as small as the geometric mean of the emittance in  $x$  and  $y$  for a flat beam, which may significantly relax the emittance requirements on diffraction limited storage rings [60, 61].

The round-to-flat transformation is a reversal of flat-to-round transformation. In this transformation, a cathode is immersed in magnetic field to generate angular-momentum dominated beam and then a triplet of skew quadrupoles

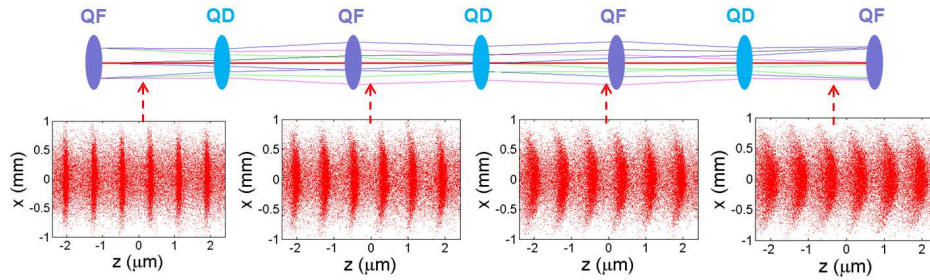
can be used to convert the round beam into a flat-beam. The round-to-flat transformation has been demonstrated at FNAL [62, 63, 64]. So far the record for the emittance ratio achieved is about 100 [64].

## 6-D MANIPULATION

6D beam manipulation typically involves several elements that couple the beam dynamics in different planes, and sometimes the dynamics is connected by second order effects.

### beam conditioning

The basic requirement of an FEL is that the electrons should slip back by one radiation wavelength as they travel one undulator period so that sustained energy exchange between the electrons and the radiation field can be maintained. However, only the reference particle can strictly satisfy this condition. Other electrons slip back by more than one wavelength because of the betatron motion. Specifically, the particles that have larger betatron amplitudes have longer path lengths. As a result, the 'wave-front' of the micro-bunches developed in FELs become curved because the particles with large betatron amplitude tend to lag behind. This effect counteracts bunching and degrades the FEL gain. This second-order smearing effect is illustrated in Fig. 9 where optical micro-bunches equally separated by 800 nm generated from laser modulation are sent through a FODO lattice. Here one can see that the micro-bunches gradually become curved in the wave-front.



**FIGURE 8.** Smearing of optical micro-bunches in a FODO lattice from beam transverse emittance.

Beam conditioning is a technique that tries to compensate for the debunching effect by establishing a correlation between particle's energy and betatron amplitude [65, 66, 67, 68, 69, 70]. Specifically, particles with larger betatron amplitudes are given extra energy to compensate for the larger path lengths such that the time-of-flight is constant for all the particles. This may greatly loosen the requirement on beam transverse emittance and enhance the gain in FELs.

Various methods have been proposed to establish such kind of correlation between particle's energy and its betatron amplitude. The first proposal uses TM210 cavities to give the beam an action-dependent energy increase [65]. Later on many proposals adopted a two-phase conditioning scheme where the electron beam is first chirped with an rf cavity; and then a focusing channel is used to delay the particles that have large betatron amplitudes; finally the beam is dechirped in an rf cavity and particles with larger betatron amplitude get an extra energy [66, 67, 68]. More exotic way to condition a beam using lasers have also been proposed: one option is to replace the two rf cavities with lasers [69]; the other is to collide the beam with high power lasers and use the radiation reaction force to condition the beam [70]. However, so far none of the methods have been tested and none of them look cheap (the main reason is that beam conditioning relies on second order effect, and therefore is not very efficient).

### Emittance partitioning

The ability to partition beam emittance among any planes is one of the ultimate goals in accelerator physics and it may open up new opportunities in accelerator science [71, 72]. For instance, the state-of-the-art photoinjector can provide a 250 pC beam with normalized emittance of  $\epsilon_{n,x}/\epsilon_{n,y}/\epsilon_{n,z}$  of 0.4/0.4/1.5  $\mu\text{m}$ . Given the same volume in 6-D phase space, if the beam emittance is partitioned to 0.1  $\mu\text{m}$  in  $x,y$  and 24  $\mu\text{m}$  in  $z$ , it will greatly improve the FEL performance. However, eigen-emittance remains constant with symplectic transformation, so it follows that there are

only two ways to partition emittance [73]: one is to affect the way the beam is born; the other is to use non-symplectic transformation and this is the only way to change eigen-emittance after the beam is born (e.g. radiation damping changes the beam eigen-emittance in  $x$  and  $y$  planes in a circular accelerator).

The round-to-flat transformation is an example where a cathode is immersed in a magnetic field to partition beam emittance in  $x$  and  $y$ . Similarly, a tilted laser may be used to illuminate the cathode to generate photoelectrons with  $x - z$  correlation, which may further allow emittance partitioning in  $x$  and  $z$  [71, 72]. One example of non-symplectic transformation is to use a tapered foil whose thickness depends on the transverse position to generate an  $x$ -dependent energy loss to the beam [74]. Analysis shows that this tapered foil can be used to partition the beam emittance in  $x$  and  $z$  [71, 75]. In principle combining the methods that affect how beam is born with the methods that provide non-symplectic maps may allow us to arbitrarily partition the emittance in eigen-planes. However, more work is needed to fully confirm the promise of these techniques and to address the technical challenges such as Coulomb scattering in the foil.

## CONCLUSIONS

In this paper, various techniques for manipulating beam distribution in 2D, 4D and 6D phase space are discussed. These techniques become a new focus of accelerator physics R&D and very likely these advanced concepts will open up new opportunities in advanced accelerators and the science enabled by them.

## ACKNOWLEDGMENTS

This work was supported by the US DOE under Contract No. DE-AC02-76SF00515.

## REFERENCES

1. R.B. Palmer, SLAC-PUB-4707 (1988).
2. B.E. Carlsten, Nucl. Instrum. Methods Phys. Res., Sect. A 285, 313 (1989).
3. L. Serafini and J.B. Rosenzweig, Phys. Rev. E 55, 7565 (1997).
4. P. Chen *et al.*, Phys. Rev. Lett. 56, 1252 (1986).
5. C. Jing *et al.*, Phys. Rev. Lett. 98, 144801 (2007).
6. R.J. England, J.B. Rosenzweig, and G. Travish, Phys. Rev. Lett. 100, 214802 (2008).
7. P. Muggli *et al.*, Phys. Rev. Lett. 101, 054801 (2008).
8. P. Piot *et al.*, Phys. Rev. Lett. 108, 034801 (2012).
9. D. Xiang *et al.*, Phys. Rev. ST Accel. Beams 15, 050707 (2012).
10. Y. Cai *et al.*, Phys. Rev. ST Accel. Beams 15, 054002 (2012).
11. P. Raimondi and A. Seryi, Phys. Rev. Lett. 86, 3779 (2001).
12. R. Ermi *et al.*, Phys. Rev. Lett. 102, 096101 (2009).
13. K. Bane *et al.*, Phys. Rev. ST Accel. Beams 12, 030704 (2009).
14. P. Emma *et al.*, Nature Photon. 4, 641 (2010).
15. Y. Ding *et al.*, Phys. Rev. Lett. 102, 254801 (2009).
16. S. Reiche *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A 593, 45 (2008).
17. P. Piot, D. Douglas, and G. Krafft, Phys. Rev. ST Accel. Beams 6, 030702 (2003).
18. C. Schroeder *et al.*, AIP Conf. Proc. 1086, 637 (2009).
19. M. Xie, Proceedings of PAC95, p183, (1995).
20. S. Reiche, LBNL CBP seminar, 2011, unpublished.
21. L.-H. Yu, Phys. Rev. A 44, 5178 (1991).
22. L.-H. Yu *et al.*, Phys. Rev. Lett. 91, 074801 (2003).
23. J. Wu and L. H. Yu, Nucl. Instrum. Methods Phys. Res., Sect. A 475, 104 (2001).
24. G. Stupakov, Phys. Rev. Lett. 102, 074801 (2009).
25. D. Xiang and G. Stupakov, Phys. Rev. ST Accel. Beams 12, 030702 (2009).
26. D. Xiang *et al.*, Phys. Rev. Lett. 105, 114801 (2010).
27. D. Xiang *et al.*, Phys. Rev. Lett. 108, 024802 (2012).
28. D. Xiang and G. Stupakov, New J. Phys. 13, 093028 (2011).
29. T. Popmintchev *et al.*, Nature Photon. 4, 822 (2010).
30. W. Kimura *et al.*, Phys. Rev. Lett. 86, 4041 (2001).
31. D. Dowell, T. Hayward and A. Vetter, Proceedings of PAC95, p992, (1995).

32. P. Emma, LCLS Technical Note, LCLS-TN-01-1, (2001).
33. J. Byrd *et al.*, Proceedings of PAC01, p380, (2001).
34. G. Stupakov and M. Zolotarev, Proceedings of FEL11, MOPB19, (2011).
35. D. Ratner and A. Chao, Proceedings of FEL11, MOPB21, (2011).
36. S. Pottorf and X.J. Wang, BNL-68013, (2000).
37. J. P. Duris, P. Musumeci, and R. K. Li, Phys. Rev. ST Accel. Beams 15, 061301 (2012).
38. D. Xiang and G. Stupakov, Phys. Rev. ST Accel. Beams 12, 080701 (2009).
39. D. Dunning *et al.*, Phys. Rev. Lett. 109, 074801 (2012).
40. G. Berden *et al.*, Phys. Rev. Lett. 99, 164801 (2007).
41. K. Bane *et al.*, Phys. Rev. ST Accel. Beams 12, 030704 (2009).
42. D. Xiang *et al.*, Phys. Rev. ST Accel. Beams 14, 112801 (2011).
43. Y. Ding *et al.*, Phys. Rev. ST Accel. Beams 14, 120701 (2011).
44. G. Andonian *et al.*, Phys. Rev. ST Accel. Beams 14, 072802 (2011).
45. D. Xiang and Y. Ding, Phys. Rev. ST Accel. Beams 13, 094001 (2010).
46. M. Cornacchia and P. Emma, Phys. Rev. ST Accel. Beams 5, 084001 (2002).
47. P. Emma *et al.*, Phys. Rev. ST Accel. Beams 9, 100702 (2006).
48. D. Xiang and A. Chao, Phys. Rev. ST Accel. Beams 14, 114001 (2011).
49. D. Xiang, Phys. Rev. ST Accel. Beams 13, 010701 (2010).
50. A. Zholents and M. Zolotarev, Argonne National Laboratory Report No. ANL/APS/LS-327, (2011).
51. Y.-E Sun *et al.*, Phys. Rev. Lett. 105, 234801 (2010).
52. P. Piot, Y.-E Sun, J.G. Power and M. Rihaoui, Phys. Rev. ST Accel. Beams 14, 022801 (2011).
53. B. Jiang *et al.*, Phys. Rev. Lett. 106, 114801 (2011).
54. S. Kowalski and H. Enge, Proceedings of the International Conference on Magnet Technology, p.182 (1972).
55. K. Hübner *et al.*, Proceedings of EPAC90, p.1297 (1990).
56. R. Talman, Phys. Rev. Lett. 74, 1590 (1995).
57. Y. Derbenev, Michigan Univ. Report UM HE 98-04, (1998).
58. Brinkmann, Y. Derbenev, and K. Fl'ttmann, Phys. Rev. ST Accel. Beams 4, 053501 (2001).
59. K.-J. Kim, Phys. Rev. ST Accel. Beams 6, 104002 (2003).
60. R. Brinkmann, Proceedings of EPAC02, p.653 (2002).
61. A. Chao and P. Raimondi, SLAC-PUB-14808 (2012).
62. D. Edwards *et al.*, Proceedings of LINAC0, p.122 (2000).
63. Y.-E Sun *et al.*, Phys. Rev. ST Accel. Beams 7, 123501 (2004).
64. P. Piot, Y.-E Sun, and K.-J. Kim, Phys. Rev. ST Accel. Beams 9, 031001 (2006).
65. A. Sessler, D. Whittum, and L.-H. Yu, Phys. Rev. Lett. 68, 309 (1992).
66. N.A. Vinokurov, Nucl. Instrum. Methods Phys. Res., Sect. A 375, 264 (1996).
67. P. Emma and G. Stupakov, Phys. Rev. ST Accel. Beams 6, 030701 (2003).
68. A. Wolski *et al.*, Phys. Rev. ST Accel. Beams 7, 080701 (2004).
69. A. Zholents, Phys. Rev. ST Accel. Beams 8, 050701 (2005).
70. C. Schroeder, E. Esarey, and W. Leemans, Phys. Rev. Lett. 93, 194801 (2004).
71. B. Carlsten *et al.*, Phys. Rev. ST Accel. Beams 14, 050706 (2011).
72. L. Duffy *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A 654, 52 (2011).
73. A. Chao, SLAC-PUB-14832 (2012).
74. J. Peterson, IEEE Trans. Nucl. Sci. 30, 2403 (1983).
75. Y. Jiao, A. Chao, and Y. Cai, SLAC-PUB-14554 (2011).