Adaptive method for electron bunch profile prediction

Alexander Scheinker[*]

Los Alamos National Laboratory, 1200 Trinity Drive, Los Alamos, New Mexico 87544, USA

Spencer Gessner[†]

SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, California 94025, USA

We report on an experiment performed at the Facility for Advanced Accelerator Experimental Tests (FACET) at SLAC National Accelerator Laboratory, in which a new adaptive control algorithm, one with known, bounded update rates, despite operating on analytically unknown cost functions, was utilized in order to provide quasi-real-time bunch property estimates of the electron beam. Multiple parameters, such as arbitrary rf phase settings and other time-varying accelerator properties, were simultaneously tuned in order to match a simulated bunch energy spectrum with a measured energy spectrum. The simple adaptive scheme was digitally implemented using Matlab and the experimental physics and industrial control system. The main result is a nonintrusive, nondestructive, real-time diagnostic scheme for prediction of bunch profiles, as well as other beam parameters, the precise control of which are important for the plasma wakefield acceleration experiments being explored at FACET.

I. INTRODUCTION

A. Motivation

The Facility for Advanced Accelerator Experimental Tests (FACET) at SLAC National Accelerator Laboratory produces high energy electron beams for plasma wakefield acceleration [1]. For these experiments, precise control of the longitudinal beam profile is very important. FACET uses an x-band transverse deflecting cavity (TCAV) to streak the beam and measure the bunch profile [Fig. 1(a)]. Although the TCAV provides an accurate measure of the bunch profile, it is a destructive measurement; the beam cannot be used for plasma wakefield acceleration (PWFA) once it has been streaked. In addition, using the TCAV to measure the bunch profile requires adjusting the optics of the final focus system to optimize the resolution and accuracy of measurement. This makes it a time consuming process and prevents on-the-fly measurements of the bunch profile during plasma experiments.

There are two diagnostics that are used as an alternative to the TCAV that provide information about the longitudinal phase space in a nondestructive manner. The first is a pyrometer that captures optical diffraction radiation (ODR) produced by the electron beam as it passes through a hole in a metal foil. The spectral content of the ODR changes with bunch length. The pyrometer is sensitive to the spectral content and the signal it collects is proportional to $1/\sigma_z$, where σ_z is the bunch length. The pyrometer is an excellent device for measuring variation in the shot-to-shot bunch profile but provides no information about the shape of the bunch profile or specific changes to shape. The second device is a nondestructive energy spectrometer consisting of a half-period vertical wiggler located in a region of large horizontal dispersion. The wiggler produces a streak of x rays with an intensity profile that is correlated with the dispersed beam profile. There x rays are intercepted by a scintillating yttrium-aluminum-garnet (SYAG) crystal and imaged by a CCD camera [Fig. 1(b)]. The horizontal profile of the x-ray streak is interpreted as the energy spectrum of the beam [2].

The measured energy spectrum is observed to correlate with the longitudinal bunch profile in a one-to-one manner if certain machine parameters, such as chicane optics, are fixed. To calculate the beam properties based on an energy spectrum measurement, the detected spectrum is compared to a simulated spectrum created with the 2D longitudinal particle tracking code, LiTrack [3]. The energy spread of short electron bunches desirable for plasma wakefield acceleration can be uniquely correlated to the beam profile if all of the various accelerator parameters which influence the bunch profile and energy spread are accounted for accurately. Unfortunately, throughout the 2 km facility, there exist systematic phase drifts of various high frequency devices, miscalibrations, and time-varying uncertainties due to thermal drifts. Therefore, in order to effectively and accurately relate an energy spectrum to a bunch profile, a very large parameter space must be searched and fit by LiTrack, which effectively limits and prevents the use of the
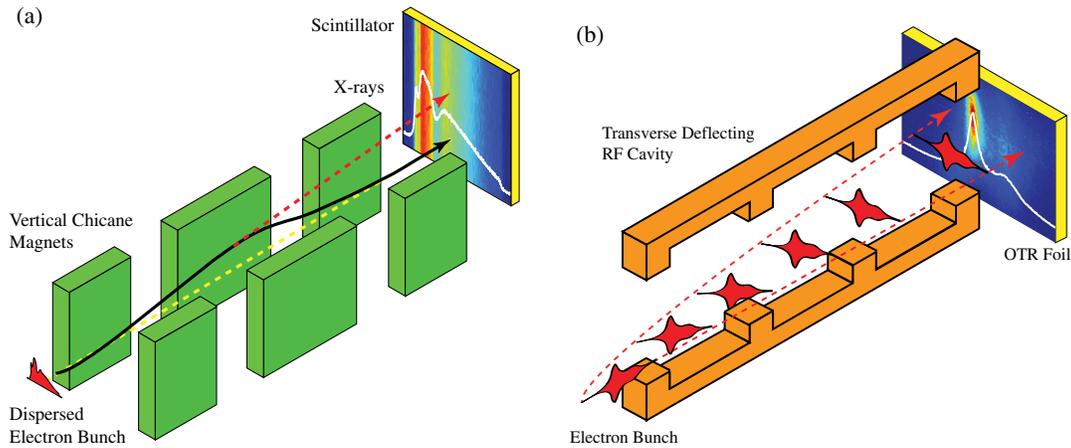
[*]ascheink@lanl.gov
[†]sgess@slac.stanford.edu

FIG. 1.   The energy spectrum is recorded as the electron bunch passes through a series of magnets and radiates x rays. The intensity distribution of the x rays is correlated to the energy spectrum of the electron beam (a). This nondestructive measurement is available at all times, and used as the input to the ES scheme, which is then matched by adaptively tuning machine parameters in the simulation. For the TCAV measurement, the electron bunch is passed through a high frequency (11.4 GHz) rf cavity with a transverse mode, in which it is streaked and passes through a metallic foil (b). The intensity of the optical transition radiation (OTR) is proportional to the longitudinal charge density distribution. This high accuracy longitudinal bunch profile measurement is a destructive technique.

energy spectrum measurement as a real-time measurement of bunch profile.

For dealing with many parameter systems, many optimization schemes [4], including in particular, genetic algorithms (GA), have been used very successfully, including the design of magnet and radio frequency (rf) cavities [5], photoinjectors [6], damping rings [7], storage ring dynamics [8], global optimization of a lattice [9], neutrino factory design [10], simultaneous optimization of beam emittance and dynamic aperture [11], free electron laser linac drivers [12] and various other accelerator physics applications [13]. The major benefit of GA-type searches is that they result in global optimization, at the cost of a lengthy search over a large range of the parameter space, and the result is only optimal relative to a known model. Robust conjugate direction search (RCDS) is another algorithm which is able to quickly optimize many parameter systems, which in addition is model independent, not requiring a detailed system knowledge, with convergence rates exceeding those of some GA approaches [14]. However, both the GA and RCDS approaches are best suited for time-invariant systems, and the goal here was to quickly track an uncertain, time-varying system. Therefore, we relied on a local, model-independent extremum seeking algorithm, whose convergence can suffer due to local minima, but whose simplicity and speed of convergence allows for a quasi-real-time tracking of a many parameter time-varying system.

B. Main result

At FACET, we coupled the above described technique with a new version of an extremum seeking (ES) algorithm, and were successful in providing a quasi-real-time estimate

of the electron bunch profile, by adaptively identifying and tracking the many uncertain, time-varying parameters required by the LiTrack code. For the ES algorithm, the cost to be minimized was the χ^2 residual between the measured and simulated energy spectra of the electron beam. System parameters such as various arbitrary phase shifts and beam properties (charge, initial phase space emittance) were the inputs to LiTrack. The adaptive scheme minimized the cost by varying an arbitrary number of parameters simultaneously. We simulate FACET while adaptively tuning multiple free parameters in a code package called LiTrackES [15,16].

C. Organization

In Sec. II we give a general outline the adaptive algorithm and how FACET can be viewed as a time-varying nonlinear system to which it is applicable. In Sec. III we describe the adaptive method and how it is applied at FACET. In Sec. IV we provide the results of the experiments conducted at FACET, demonstrating an ability to predict accelerator component settings and electron beam properties. In the Appendix we provide a detailed theoretical statement of the convergence properties of the adaptive algorithm.

II. APPLICATION TO PREDICTING FACET BEAM PROPERTIES

ES, as described here, was recently developed for the stabilization of uncertain, open-loop unstable systems [17,18]. The mechanism behind this new form of ES, that of introducing high frequency oscillations into a system's dynamics, is closely related to the field of vibrational control, such as stabilizing the vertical equilibrium point of

a pendulum by quickly oscillating its pivot point [19,20]. Recently, a new, bounded form of ES has been developed [21], one with analytically guaranteed update rates and control efforts, and implemented in simulation [22] and demonstrated in hardware [23] to tune multiple components of particle accelerators. Although operating on analytically unknown functions, limits are guaranteed by the unknown cost function entering the scheme's dynamics as the argument of a known, bounded oscillatory function, such as sine or cosine. A detailed, analytic description of the algorithm is presented in the Appendix.

In this paper, in applying this new form of ES to the parameter estimation problem [24], we consider FACET as a system of the form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{p}, t), \tag{1}$$

where $\mathbf{f}(\mathbf{x}, \mathbf{p}, t)$ represents the actual, analytically uncertain dynamics of the accelerator, $\mathbf{x}(t) = [x_1(t), ..., x_n(t)]$ represents various aspects of the beam, such as bunch length, beam energy, bunch charge, etc. at certain locations throughout the accelerator, and $\mathbf{p}(t) = [p_1(t), ..., p_n(t)]$ represents various time-varying uncertain parameters of the accelerator itself, such as rf system phase drifts and rf field amplitudes throughout the machine.

We consider, relative to (1), the approximation of FACET given by LiTrack:

$$\dot{\hat{\mathbf{x}}} = \hat{\mathbf{f}}(\hat{\mathbf{x}}, \hat{\mathbf{p}}, t), \tag{2}$$

where $\hat{\mathbf{f}}(\hat{\mathbf{x}}, \hat{\mathbf{p}}, t)$ represents the approximated simulation of the system's dynamics, where actual beam properties $\mathbf{x}(t)$ are approximated by their simulated estimates $\hat{\mathbf{x}}(t) = [\hat{x}_1(t), ..., \hat{x}_n(t)]$ and actual system parameters, $\mathbf{p}(t)$, are approximated by virtual parameters $\hat{\mathbf{p}}(t) = [\hat{p}_1(t), ..., \hat{p}_n(t)]$.

Our goal is to closely approximate certain actual beam properties, $x_i(t)$, such as the longitudinal beam profile, by their virtual estimates $\hat{x}_i(t)$, which is only possible if we closely approximate the true machine parameters $\mathbf{p}(t)$, by adaptively tuning the virtual parameters $\hat{\mathbf{p}}(t)$.

We perform this adaptation based on a cost function, C, whose values depend on the comparison between a detected, analytically uncertain and noisy measurement

$$\tilde{y} = h(\mathbf{x}, \mathbf{p}, t) + n(t) \tag{3}$$

and a simulation of that same measurement

$$\hat{y} = \hat{h}(\hat{\mathbf{x}}, \hat{\mathbf{p}}, t). \tag{4}$$

The adaptive tuning law is

$$\dot{\hat{p}}_i = \sqrt{\alpha_i \omega_i} \cos\{\omega_i t + k_i C[\tilde{y}(\mathbf{x}, \mathbf{p}, t), \hat{y}(\hat{\mathbf{x}}, \hat{\mathbf{p}}, t)]\}, \tag{5}$$

which, according to Theorem 1, has average dynamics

$$\dot{\bar{\hat{p}}}_i = -\frac{k_i \alpha_i}{2} \frac{\partial C}{\partial \bar{\hat{p}}_i}, \tag{6}$$

performing a gradient descent to locally minimizing values $\hat{p}_i(t)$ of $C(t)$, where the goal is to match the actual accelerator parameters $\mathbf{p}(t)$, and to match the output of the simulation, $\hat{\mathbf{x}}(t)$, to the actual electron beam $\mathbf{x}(t)$.

The uniqueness of this convergence heavily depends on the actual analytic forms of C, h, and $\mathbf{f}$, such a convergence may not always provide an accurate or unique prediction of actual beam $x_i(t)$ and machine $p_i(t)$ parameter values based on their virtual observers $\hat{x}_i(t)$ and $\hat{p}_i(t)$. In this case, as will be shown below, we are able to closely track the actual beam characteristics.

III. ADAPTIVE TUNING METHOD AT FACET

Figures 1–4 show the beam measurement devices and overall setup of the tuning procedure at FACET. Figure 1 shows the setup of the destructive TCAV bunch profile measurement and nondestructive SYAG energy spread measurement. Figure 2 shows actual recorded bunch profile and energy spread spectrums. In Fig. 3 an overview of the coupled simulator-accelerator setup is shown. A simulation of the accelerator, LiTrack is run in parallel to the machines operation. The simulation was initialized with guesses and any available measurements of actual machine settings, $\mathbf{p} = (p_1, ..., p_n)$. We emphasize that these are only guesses because even measured values are noisy and have arbitrary phase shift errors. The electron beam in the actual machine was accelerated and then passed through a series of deflecting magnets, as shown in Figs. 1(b) and 2, which created x rays, whose intensity distribution can be correlated to the electron bunch density via LiTrack. This nondestructive measurement is available at all times, and used as the input to the ES scheme, which is then matched by adaptively tuning machine parameters in the simulation. Once the simulated and actual spectrum were matched, certain beam properties could be predicted by the simulation. Figure 4 shows the result of one such adaptation, as the cost function is minimized the LiTrackES spectrum (green) converges to the actual spectrum detected on the SYAG (black).

Each parameter setting has its own influence on electron beam dynamics, which in turn influenced the separation, charge, length, etc. of the leading and trailing electron bunches.

The cost that our adaptive scheme was attempting to minimize was then the difference between the actual, detected spectrum, and that predicted by LiTrack:

$$C(\mathbf{x}, \hat{\mathbf{x}}, \mathbf{p}, \hat{\mathbf{p}}, t) = \int |\tilde{\psi}(\mathbf{x}, \mathbf{p}, t, \nu) - \hat{\psi}(\hat{\mathbf{x}}, \hat{\mathbf{p}}, t, \nu)|^2 d\nu, \tag{7}$$

in which $\tilde{\psi}(\mathbf{x}, \mathbf{p}, t, \nu)$ was a noisy measurement of the actual, time-varying (due to phase drift, thermal cycling, ...) energy spectrum, and $\hat{\psi}(\hat{\mathbf{x}}, \hat{\mathbf{p}}, t, \nu)$ was the LiTrack, simulated spectrum, $\mathbf{x}(t) = [x_1(t), ..., x_n(t)]$ represents various
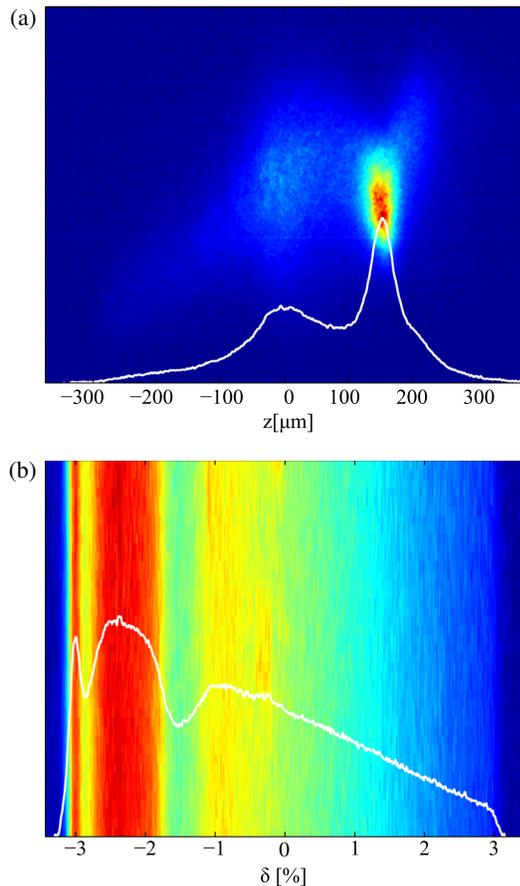
FIG. 2.   Typical bunch profile measurement (a). Typical energy spectrum measurement (b). Color, from blue to red, indicates increasing detected intensity. The white curves are created by integrating the detected intensities vertically along the detectors, which gives a measure of longitudinal beam profile (a), and a measure of deviation from average energy (b).

aspects of the beam, such as bunch length, beam energy, bunch charge, etc. at certain locations throughout the accelerator, $\mathbf{p}(t) = [p_1(t), \dots, p_n(t)]$ represents various time-varying uncertain parameters of the accelerator itself, such as rf system phase drifts and rf field amplitudes throughout the machine, $\mathbf{x}(t)$ are approximated by their simulated estimates $\hat{\mathbf{x}}(t) = [\hat{x}_1(t), \dots, \hat{x}_n(t)]$ and actual system parameters, $\mathbf{p}(t)$, are approximated by virtual parameters $\hat{\mathbf{p}}(t) = [\hat{p}_1(t), \dots, \hat{p}_n(t)]$.

The problem was then to minimize the measurable, but analytically unknown function $C(\mathbf{x}, \hat{\mathbf{x}}, \mathbf{p}, \hat{\mathbf{p}}, t) : \mathbb{R}^{4n} \times \mathbb{R}^+ \to \mathbb{R}$, by adaptively tuning the simulation parameters $\hat{\mathbf{p}}$. The hope was that, by finding simulation machine settings which resulted in matched spectrums, we would also match other properties of the real and simulated beams, something we could not simply do by setting the simulation parameters to the exact machine settings, due to unknowns, such as time-varying, arbitrary phase shifts.

The first step of the adaptive scheme was to choose physically realizable constraints for all parameters: $\mathbf{p}_{\max} = (p_{1,\max}, \dots, p_{m,\max})$, $\mathbf{p}_{\min} = (p_{1,\min}, \dots, p_{m,\min})$. Implementing initial parameter settings $\mathbf{p}(1)$, which are chosen based on the physics model, experience, and available measurements, allowed us to measure $C(1, t_1) = C[\mathbf{x}(1), \hat{\mathbf{x}}(1), \mathbf{p}(1), \hat{\mathbf{p}}(1), t_1]$. The iterative update scheme was then

$$p_i(n+1) = p_i(n) + \Delta\sqrt{\alpha\omega_i}\cos[\omega_i n\Delta + kC(n, t_n)], \quad (8)$$

where t_n represents actual time at which parameter updates are made, something limited by sampling rates and LiTrackES simulation speed. In our application, the beam repetition rate was ~ 1 Hz, and $dt = t_{n+1} - t_n \approx 0.6s$.

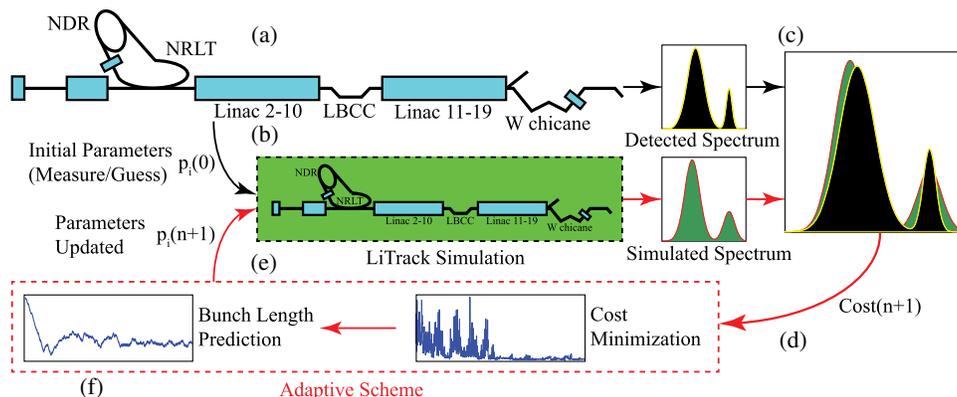Update law (8) is based on the finite difference approximation of the derivative:

FIG. 3.   The actual FACET accelerator (a) is run along with the LiTrack simulator (b), whose initial parameter inputs, $p_i(0)$, are a combination of accurately measured settings and guesses for time varying, not easily detectable actual settings, such as drifting rf phase. The measured and predicted energy spectrums are compared (c) and a cost, $C(n+1)$ is calculated based on the mismatch. The cost is fed into the adaptive scheme (d), as described above, and the parameters are automatically tuned and updated (e). When the cost is oscillating near minimum, the bunch length prediction is most accurate (f).
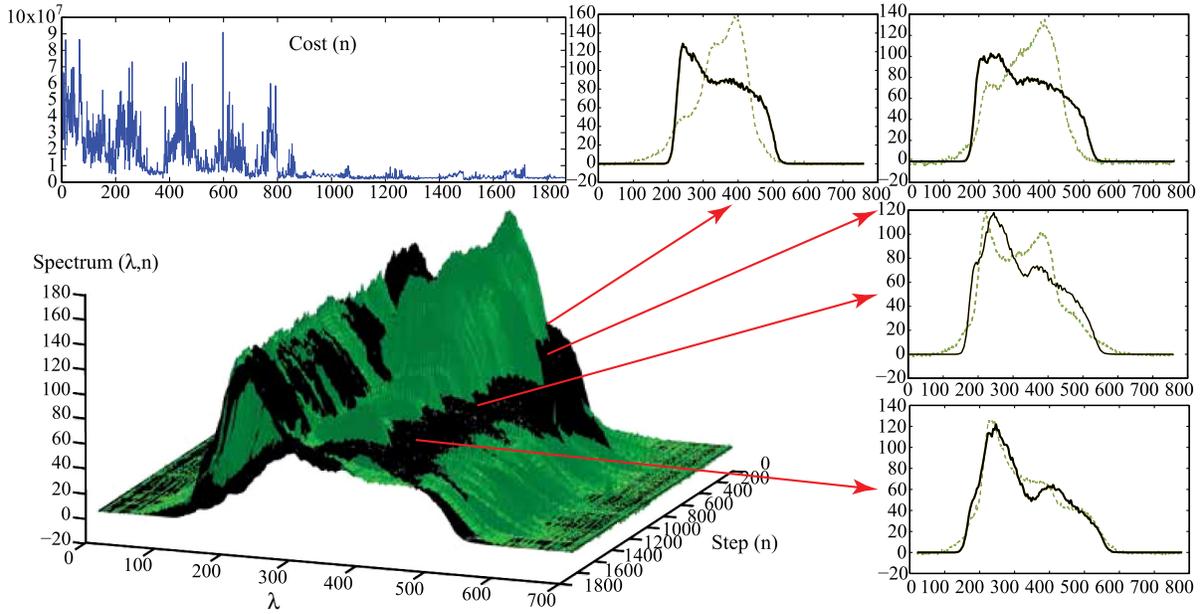
FIG. 4.   Convergence of the LiTrack spectrum to the actual, time-varying spectrum for a fixed rms-width beam.

$$\frac{p_i(t+\Delta)-p_i(t)}{\Delta} \approx \frac{\partial p_i}{\partial t} = \sqrt{\alpha\omega_i}\cos\{\omega_i t + kC[\mathbf{p}(t),t]\},$$

$$(9)$$

which we may expand as

$$\cos(\omega_i t + kC) = \cos(\omega_i t)\cos(kC) - \sin(\omega_i t)\sin(kC) \quad (10)$$

and rewrite the p_i $(1 \leq i \leq n)$ dynamics as

$$\dot{p}_i = \sqrt{\omega_i}\cos(\omega_i t)\sqrt{\alpha}\cos(kC) - \sqrt{\omega_i}\sin(\omega_i t)\sqrt{\alpha}\sin(kC),$$

$$(11)$$

which, according to Theorem 1, resulted in an average parameter and cost relationship of the form

$$\dot{\bar{p}}_i = -\frac{k\alpha}{2}\frac{\partial C(\bar{\mathbf{p}},t)}{\partial \bar{p}_i}\{\cos^2[kC(\bar{\mathbf{p}},t)] + \sin^2[kC(\bar{\mathbf{p}},t)]\}$$

$$= -\frac{k\alpha}{2}\frac{\partial C(\bar{\mathbf{p}},t)}{\partial \bar{p}_i}, \quad\quad (12)$$

and therefore

$$\dot{\bar{\mathbf{p}}} = -\frac{k\alpha}{2}\nabla C, \quad\quad (13)$$

which is a gradient descent towards the minimum of C. The constraints were simply implemented by checking the updated parameters at each step and confining them to their bounds if necessary. Because the values of different parameters p_i differed by orders of magnitude, they each required individual values of k_i and α_i.

The ES algorithm is constantly updating the virtual machine parameters to provide a better match between the simulated and energy spectrum. It outputs a new estimate of the longitudinal bunch profile for every iteration of the loop.

IV. RESULTS

LiTrackES simulates large components of FACET as single elements. The critical elements of the simulation are the north damping ring (NDR) which sets the initial bunch parameters, including the bunch length and energy spread, the north ring to linac (NRTL) which is the first of three bunch compressors, linac sectors 2–10 where the beam is accelerated and chirped, the second bunch compressor in sector 10 (LBCC), linac sectors 11–19 where the beam is again accelerated and chirped, and finally the FACET W-chicane which is the third and final bunch compressor.

We calibrated the LiTrackES algorithm using simultaneous measurements of the energy spectrum and bunch profile while allowing a set of unknown parameters to converge. After convergence we left a subset of these calibrated parameters fixed, as they are known to vary slowly or not at all and performed our tuning on a much smaller subset of the parameters: (i) p_1: NDR bunch length; (ii) p_2: NRTL energy offset; (iii) p_3: NRTL compressor amplitude; (iv) p_4: NRTL chicane T_{566}; and (v) p_5: phase ramp. "Phase ramp" refers to a net phase of the NDR and NRTL rf systems with respect to the main linac rf. Changing the phase ramp parameter results in a phase set offset in the linac relative to some desired phase.

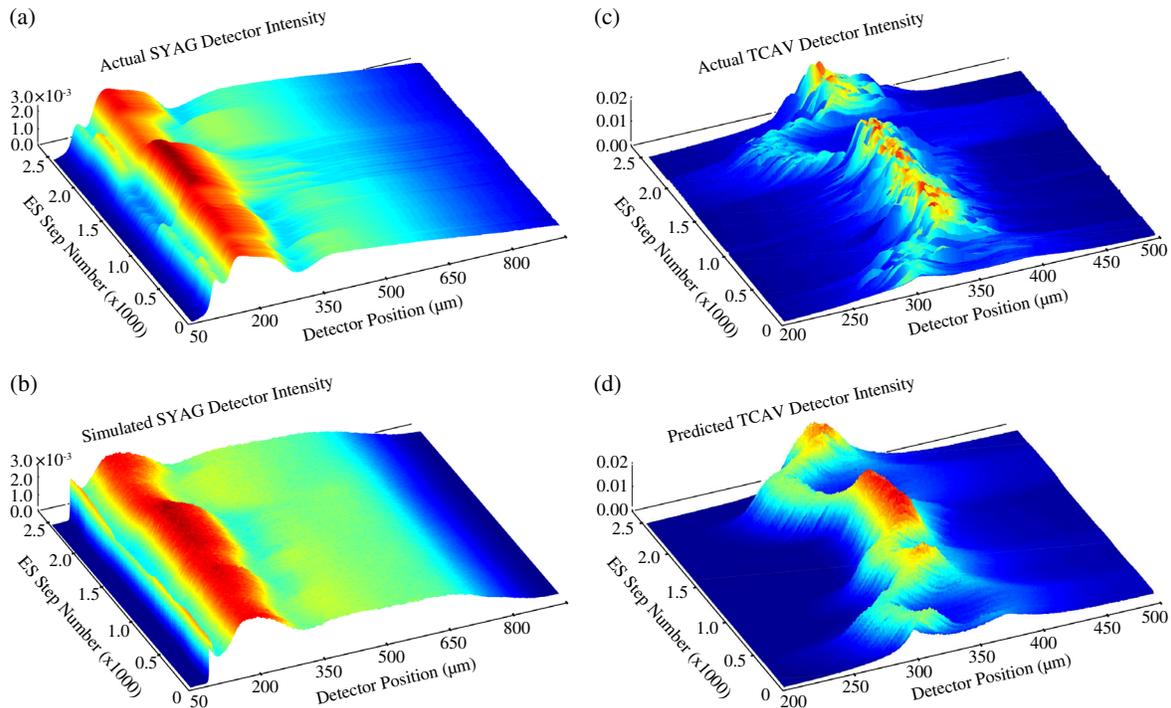In this experiment, we had the added benefit of continuously observing the bunch profile using the TCAV,

FIG. 5. (a) The energy spectrum, which is always available as a nondestructive measurement, was detected and imported into LiTrackES in order to calculate a cost on which the adaptive convergence was based. (b) LiTrackES was able to adaptively tune parameters $\hat{p}_i(t)$ in order to track the detected energy spectrum with the simulated spectrum. (c) The TCAV was on and recording the longitudinal beam profile during the adaptive process, these values were not used as inputs to LiTrackES, but were detected in order to measure the predictive performance of LiTrackES. (d) Based only on energy spectrum readings, LiTrackES was able to predict the longitudinal electron bunch profile, tracking the actual bunch profile measurements with the LiTrackES output. Color, from blue to red, indicates increasing detected intensity.

while predicting it based on a match with the measured energy spectrum, so that we could have a more detailed comparison of the predicted and actual beam characteristics. Our approach was to periodically vary FACET's phase ramp, which resulted in a significant change in the longitudinal bunch profile. We then allowed LiTrackES to adapt to the changing energy spectrum and thereby attempt to predict the new beam profile. The results of the experiment are shown in Figs. 5–9 and discussed in detail below.

A. Phase ramp prediction

The phase ramp parameter is used to tune out correlated phase errors that accumulate throughout the linac. As an example, phase ramp is tuned diurnally to offset the contraction and expansion of the main drive line which provides the reference rf for all systems in the linac. Adjustments to phase ramp occur a few times per day and the beam is allowed to settle after the changes are made. The phase drifts that occur between changes to the phase ramp parameter are not measured. Instead, certain beam features, such as the energy spectrum and bunch profile will change noticeably. If they deviate too much from their desired state, an adjustment to phase ramp is made to bring them back. Tracking of phase ramp changes

is demonstrated in Fig. 6. As seen in the figure, the time-varying phase was a moving trajectory that the ES algorithm was attempting to track. Although its initial velocity was incorrect, the prediction is seen to quickly begin following, catch up to, and then remain locked within close proximity of the time-varying measured value.

B. Impact of phase ramp on bunch profile

Phase ramp is one of the few parameters that are regularly tuned to achieve the desired "two-bunch" profiles for PWFA experiments. This is because phase ramp is
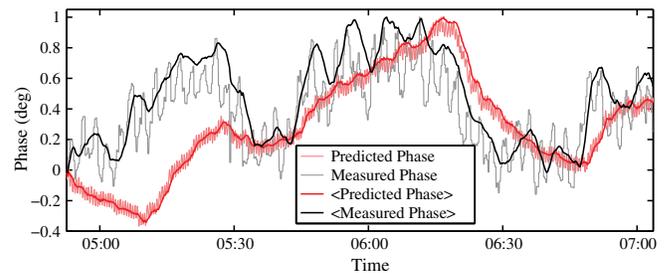

FIG. 6. Estimation of the rf phase in the first ten sectors of the linac using LiTrackES compared to the vector sum of klystron phases.
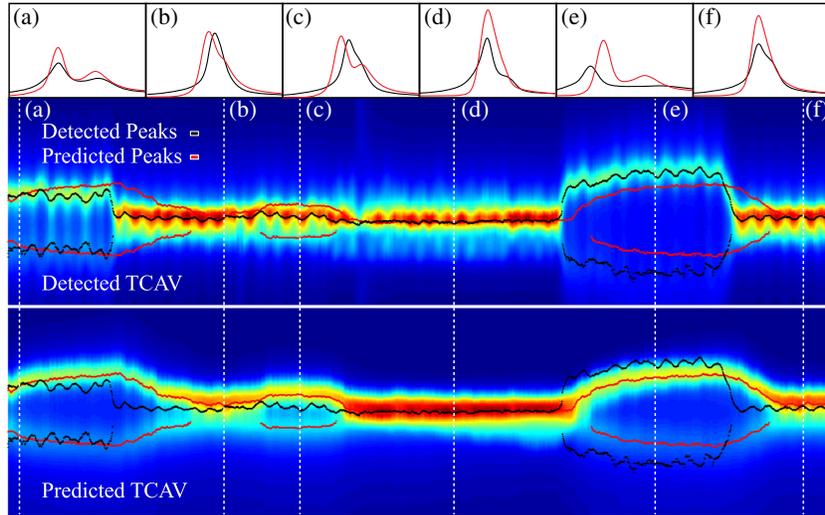
FIG. 7.   A top-down view of the actual and predicted bunch profile measurements shows that LiTrackES was able to track changes in the bunch profile, including a successful prediction of when there was a single bunch and when it was split in two as is done in driver-witness plasma wakefield acceleration experiments. Color, from blue to red, indicates increasing detected intensity. The red and black lines placed on top of the detector intensities show the predicted and measured peaks, respectively. Plots (a)–(f) show actual and predicted longitudinal beam profiles as measured and predicted along the white dashed lines (a)–(f).
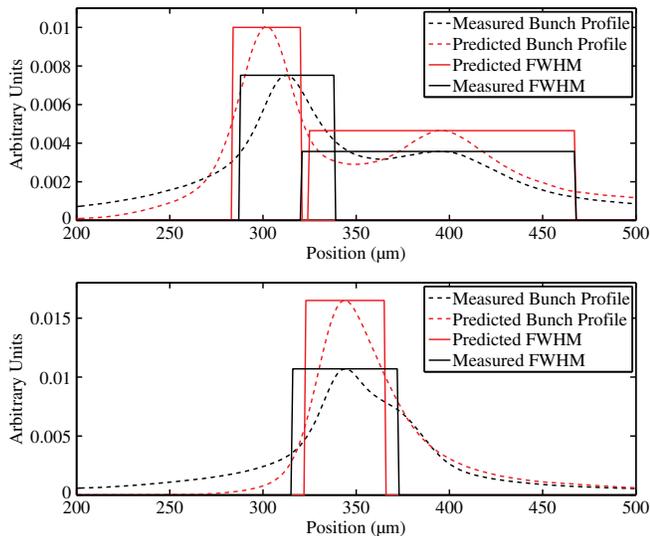

FIG. 8.   Moving averages of the predicted and actual bunch profiles were used in order to estimate bunch width with a measurement of the full width half maximum.

directly tied to the injection phase of the linac. The beam is injected 21° ahead of crest in the first ten sectors of the linac. The beam is both accelerated and chirped in this linac section and then compressed by a magnetic chicane at the end of sector 10. The extent to which the bunch is compressed depends sensitively on the beam chirp, and therefor the phase in the first ten linac sectors. For this reason, a feedback that can actively determine phase drift in the linac would benefit PWFA experiments at FACET. The LiTrackES algorithm demonstrates this ability as shown in Fig. 6.

C. Bunch profile prediction

Experiments performed at FACET utilize a wide array of bunch profiles, including separate drive-witness bunches in some of the plasma wakefield acceleration experiments. We have confirmed that LiTrackES is able to quickly provide an accurate, nondestructive bunch profile measurement prediction even as machine and initial beam conditions slowly
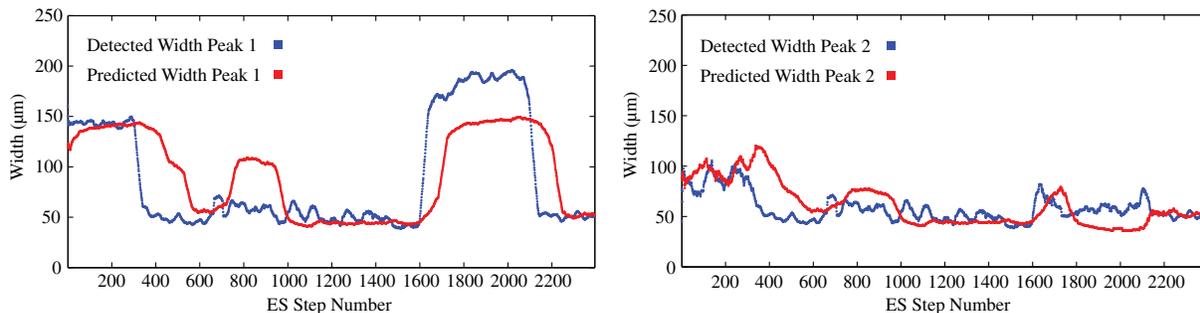

FIG. 9.   While predicting the beam profile, an estimate of the full width half maximums of the bunch peaks was made, which sometimes coincided when the beam was a single bunch.

vary with time, and can distinguish between single and double bunch settings, as shown in Figs. 8 and 9. Any experiments performed at FACET for which bunch profile information is important can now utilize LiTrackES, running the program continuously in the background, recording bunch profile predictions to later be correlated with experimental results. In particular, by measuring the charge and bunch length of the drive and witness beams in a nondestructive manner, we are able to determine properties of the plasma wake that would otherwise be inaccessible, specifically the transformer ratio and other beam-loading related quantities.

V. CONCLUSIONS AND FUTURE WORK

LiTrackES, the combination of ES and LiTrack, as demonstrated, is able to provide a quasi-real-time estimate of many machine and electron beam properties which are either inaccessible or require destructive measurements. We plan to improve the convergence rate of LiTrackES by fine-tuning the adaptive scheme's parameters, such as the gains k_i, perturbing amplitudes α_i and dithering frequencies ω_i. Furthermore, we plan on taking advantage of several simultaneously running LiTrackES schemes, which can communicate with each other in an intelligent way, and each of which has slightly different adaptive parameters/ initial parameter guesses, which we believe can greatly increase both the rate and accuracy of the convergence. Another major goal is the extension of this algorithm from monitoring to tuning. Preliminary work has begun on a user-friendly graphical user interface version of LiTrackES, so that machine operators can take advantage of various arbitrary phase predictions, which may be useful for tuning. Furthermore, we hope to one day utilize LiTrackES as an actual feedback to the machine settings in order to tune for desired electron beam properties.

ACKNOWLEDGMENTS

APPENDIX: ALGORITHM THEORETICAL BACKGROUND

The ES scheme is applicable to systems of differential equations of the form

$$\dot{x}_i = f_i(\mathbf{x}, t) + g_i(\mathbf{x}, t) u_i(\mathbf{x}, t), \qquad (A1)$$

where $\dot{h} \equiv \frac{\partial h}{\partial t}$, in which $\mathbf{f}(\mathbf{x}, t) = [f_1(\mathbf{x}, t), ..., f_n(\mathbf{x}, t)]$ and $\mathbf{g}(\mathbf{x}, t) = [g_1(\mathbf{x}, t), ..., g_n(\mathbf{x}, t)]$ are unknown or uncertain functions and $\mathbf{u}(\mathbf{x}, t) = [u_1(\mathbf{x}, t), ..., u_n(\mathbf{x}, t)]$ are control inputs of our choice. Our goal is to minimize an analytically unknown, but available for measurement time-varying

"cost function," $V(\mathbf{x}, t)$, which depends on multiple parameters, $\mathbf{x} = [x_1(t), ..., x_n(t)]$, whose measurement may be noise corrupted, so that the actual available measurements are of the form $\tilde{V}(\mathbf{x}, t) = V(\mathbf{x}, t) + n(t)$, in which $n(t)$ is uncertain, additive measurement noise.

According to the results of [17,18,22], choosing controllers of the form

$$u_i(\mathbf{x}, t) = \sqrt{\alpha \omega_i} \cos[\omega_i t + k\tilde{V}(\mathbf{x}, t)], \qquad (A2)$$

results in an average gradient descent of the function $V(\mathbf{x}, t)$ of the form

$$\dot{\bar{x}}_i = f_i(\bar{\mathbf{x}}, t) - \frac{k\alpha}{2} g_i^2(\bar{\mathbf{x}}, t) \frac{\partial V(\bar{\mathbf{x}}, t)}{\partial \bar{x}_i}, \quad \bar{x}_i(0) = x_i(0). \quad (A3)$$

Remark 1.—Several important benefits of this controller are as follows.

(i) The minimization of $V(\bar{\mathbf{x}}, t)$, according to (A3), is achieved by choosing a sufficiently large value of the factor $k\alpha > 0$, relative to an upper bound on $|f_i(\mathbf{x}, t)|$ and a lower bound on $|g_i^2(\mathbf{x}, t)|$, without knowing their analytic forms.

(ii) We do not need to know the sign of the unknown functions $g_i(\mathbf{x}, t)$ and they may change their signs as a function of time, that is, we do not need to know in which direction our control inputs are acting, whether they are pulling or pushing, because $g_i^2(\mathbf{x}, t) \geq 0$.

(iii) Although we are only able to measure the noise-corrupted signal $\tilde{V}(\bar{\mathbf{x}}, t) = V(\bar{\mathbf{x}}, t) + n(t)$, on average the gradient descent takes place relative to the actual function of interest $V(\bar{\mathbf{x}}, t)$, as in (A3).

Remark 2.—The $\cos(\cdot)$ terms in (A2) may be replaced by $\sin(\cdot)$ functions, or the two can be mixed together. The only requirement, for the convergence of the scheme, is that the perturbing functions are orthogonal in the frequency domain, such as $\sin(\omega_i)$ and $\sin(\omega_j)$, where $\omega_i = \omega r_i$ and $r_i \neq r_j$, or $\cos(\omega_i)$ and $\sin(\omega_i)$. By orthogonality in the frequency domain, we mean that the products of the functions uniformly converge to zero in the L^2 norm on compact sets of the form $[0, T]$, as ω is increased to infinity. For example, considering $\sin(\omega_1 t) = \sin(\omega r_1 t)$, $\sin(\omega_2 t) = \sin(\omega r_2 t)$, the limit as $\omega \to \infty$ of the inner product in $L^2[0, T]$, for any $T > 0$ is

$$\lim_{\omega \to \infty} \langle \sin(\omega_1 t), \sin(\omega_2 t) \rangle = \lim_{\omega \to \infty} \int_0^T \sin(\omega r_1 t) \sin(\omega r_2 t) dt$$
$$= 0.$$

Remark 3.—The terms, k, α, and ω_i in the controller (A2) have varying influence on the overall convergence.

(i) The term k can be thought of as the control gain. Increasing k artificially inflates the value of the cost function, via the product $kV(\bar{\mathbf{x}}, t)$ and its gradient, resulting in faster convergence. Furthermore, if desired, such as for

parameters of varying orders of magnitude, each parameter may implement its own unique k_i.

(ii) The term α plays two roles. First, as another factor of the overall gain because it enters the average dynamics via the product $-\frac{k\alpha}{2}\frac{\partial V(\bar{\mathbf{x}},t)}{\partial \bar{x}_i}$. Also, the term α plays the role of the dithering amplitude, which is the extent to which a parameter that we are adapting will oscillate and therefore its ability to search for a minimum and to escape local minimums, both of which are improved with increased α. Increasing α too much may be problematic because at steady state, when $\frac{\partial V(\bar{\mathbf{x}},t)}{\partial \bar{x}_i} \approx 0$, the perturbed parameters will perform oscillations about their optimal values of magnitude $\sqrt{\frac{\alpha}{\omega}}$. Furthermore, if desired, such as for parameters of varying orders of magnitude, each parameter may implement its own unique α_i.

(iii) The terms ω_i are the frequencies at which parameters are dithered. In order for the scheme to be effective, and for the averaging analysis to hold, the ω_i must be sufficiently large compared to all other system dynamics and relative to the values of k and α. Once a sufficiently large ω_i is chosen, if k or α is increased too much, the scheme may become unstable unless ω_i is increased sufficiently. Furthermore, each parameter must have a different ω_i so that they all evolve independently, performing gradient descents relative to their own partial derivatives.

For completeness, we present the complete statement of the general averaging result for highly oscillatory, uncertain systems of differential equations, which we have described above. The proof and more details regarding the adaptive scheme are available in [17,18,22].

Theorem 1 [17,18,22].—Consider the following nonlinear system of differential equations:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},t) + \sum_{i=1}^{n} \mathbf{e}_i g_i(\mathbf{x},t)\sqrt{\omega_i}\cos[\omega_i t + k\tilde{V}(\mathbf{x},t)], \quad (A4)$$

where $\omega_i \neq \omega_j$, $\forall i \neq j$, the functions

$$\mathbf{f}(\mathbf{x},t)\colon \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n \quad (A5)$$

$$g_i(\mathbf{x},t)\colon \mathbb{R}^n \times \mathbb{R} \to \mathbb{R} \quad (A6)$$

$$\tilde{V}(\mathbf{x},t)\colon \mathbb{R}^n \times \mathbb{R} \to \mathbb{R} \quad (A7)$$

are continuously differentiable with respect to $\mathbf{x}$, piecewise continuous with respect to t, and the functions $\mathbf{e}_i$ are the ith basis vectors of $\mathbb{R}^n$. The components $x_i(t)$ of $\mathbf{x}(t)$ satisfy the dynamics

$$\dot{x}_i = f_i(\mathbf{x},t) + g_i(\mathbf{x},t)\sqrt{\omega_i}\cos[\omega_i t + k\tilde{V}(\mathbf{x},t)]. \quad (A8)$$

The function $\tilde{V}(\mathbf{x},t) = V(\mathbf{x},t) + n(t)$, is a noise-corrupted measurement of the analytically unknown function $V(\mathbf{x},t)$.

Relative to system (A4), we consider the following average system of differential equations:

$$\dot{\bar{\mathbf{x}}} = \mathbf{f}(\bar{\mathbf{x}},t) - \frac{k\alpha}{2}\mathbf{g}(\bar{\mathbf{x}},t)\mathbf{g}^T(\bar{\mathbf{x}},t)[\boldsymbol{\nabla}V(\bar{\mathbf{x}},t)]^T, \quad (A9)$$

where $\bar{\mathbf{x}}(0) = \mathbf{x}(0)$ and $\mathbf{g} = (g_1, ..., g_n)$. The components $\bar{x}_i(t)$ of the trajectory $\bar{\mathbf{x}}(t)$ of (A9) satisfy the dynamics

$$\dot{\bar{x}}_i = f_i(\bar{\mathbf{x}},t) - \frac{k\alpha}{2}g_i^2(\mathbf{x},t)\frac{\partial V(\bar{\mathbf{x}},t)}{\partial \bar{x}_i}, \quad \bar{x}_i(0) = x_i(0). \quad (A10)$$

For any $T > 0$, any compact set $K \subset \mathbb{R}^n$, and any $\delta > 0$, there exists $\omega^\star$ such that for all $\omega > \omega^\star$, the distance between the trajectory $\mathbf{x}(t)$ of system (A4) and the average system trajectory $\bar{\mathbf{x}}(t)$ of system (A9), satisfy the bounds

$$\max_{\mathbf{x},\bar{\mathbf{x}}\in K, t\in[0,T]} \|\mathbf{x}(t) - \bar{\mathbf{x}}(t)\| < \delta. \quad (A11)$$

Furthermore, if a trajectory, $\mathbf{x}^\star(t)$, of the average system (A9) is asymptotically stable, then the trajectory of the actual system (A4) approaches and remains within δ of that trajectory as well, that is

$$\lim_{t\to\infty}\|\bar{\mathbf{x}}(t) - \mathbf{x}^\star(t)\| = 0 \Rightarrow \lim_{t\to\infty}\|\mathbf{x}(t) - \mathbf{x}^\star(t)\| < \delta. \quad (A12)$$

[1] M. J. Hogan, T. O. Raubenheimer, A. Seryi, P. Muggli, T. Katsouleas, C. Huang, W. Lu, W. An, K. A. Marsh, W. B. Mori, C. E. Clayton, and C. Joshi, Plasma wakefield acceleration experiments at FACET, New J. Phys. **12**, 055030 (2010).

[2] J. Seeman, W. Brunk, R. Early, M. Ross, E. Tillman, and D. Walz, SLC energy spectrum monitor using synchrotron radiation, Report No. SLAC-PUB-3495, 1986.

[3] K. Bane and P. Emma, LiTrack: A fast longitudinal phase space tracking code, Report No. SLAC-PUB-11035, 2005.

[4] M. Borland, Report No. APS LS-287, 2000.

[5] R. Hajima, N. Taked, H. Ohashi, and M. Akiyama, Optimization of wiggler magnets ordering using a genetic algorithm, Nucl. Instrum. Methods Phys. Res., Sect. A **318**, 822 (1992).

[6] I. Bazarov and C. Sinclair, Multivariate optimization of a high brightness dc gun photo injector, Phys. Rev. ST Accel. Beams **8**, 034202 (2005).

[7] L. Emery, in Proceedings of the 21st Particle Accelerator Conference, Knoxville, 2005 (IEEE, Piscataway, NJ, 2005).

[8] M. Borland, V. Sajaev, L. Emery, and A. Xiao, in Proceedings of the 23rd Particle Accelerator Conference, Vancouver, Canada, 2009 (IEEE, Piscataway, NJ, 2009).

[9] L. Yang, D. Robin, F. Sannibale, C. Steier, and W. Wan, Global optimization of an accelerator lattice using multiobjective genetic algorithms, Nucl. Instrum. Methods Phys. Res., Sect. A **609**, 50 (2009).

ALEXANDER SCHEINKER AND SPENCER GESSNER

Phys. Rev. ST Accel. Beams **18,** 102801 (2015)

[10] A. Poklonskiy and D. Neuffer, Evolutionary algorithm for the neutrino factory front end design, Int. J. Mod. Phys. A **24**, 959 (2009).

[11] W. Gao, L. Wang, and W. Li, Simultaneous optimization of beam emittance and dynamic aperture for electron storage ring using genetic algorithm, Phys. Rev. ST Accel. Beams **14**, 094001 (2011).

[12] R. Bartolini, M. Apollonio, and I. P. S. Martin, Multi-objective genetic algorithm optimization of the beam dynamics in linac drivers for free electron lasers, Phys. Rev. ST Accel. Beams **15**, 030701 (2012).

[13] A. Hofler, B. Terzic, M. Kramer, A. Zvezdin, V. Morozov, Y. Roblin, F. Lin, and C. Jarvis, Innovative applications of genetic algorithms to problems in accelerator physics, Phys. Rev. ST Accel. Beams **16**, 010101 (2013).

[14] X. Huang, J. Corbett, J. Safranek, and J. Wu, An algorithm for online optimization of accelerators, Nucl. Instrum. Methods Phys. Res., Sect. A **726**, 77 (2013).

[15] S. J. Gessner, E. Adili, F. J. Decker, M. J. Hogan, T. Raubenheimer, and A. Scheinker, Longitudinal phase space dynamics with novel diagnostic techniques at FACET, *Proceedings of the International Particle Accelerator Conference in Shanghai China, 2013* (2013).

[16] https://github.com/aard-pwfa/LiT_ES_Scan.

[17] A. Scheinker, Extremum seeking for stabilization, Ph.D. thesis, University of California, San Diego, 2012. http:// www.alexscheinker.com/Alexander_Scheinker_Thesis_w _Magnets.pdf.

[18] A. Scheinker and M. Krstic, Minimum-seeking for CLFs: Universal semiglobally stabilizing feedback under unknown control directions, IEEE Trans. Autom. Control **58**, 1107 (2013).

[19] P. L. Kapitza, Dynamic stability of the pendulum with vibrating suspension point, *Sov. Phys. JETP* **21**, *588 (1951)*; P. L. Kapitza, in *Collected Papers of P. L. Kapitza*, edited by T. Haar (Pergamon, London, 1965), Vol. 2, pp. 714–726.

[20] S. Meerkov, Principle of vibrational control: Theory and applications, IEEE Trans. Autom. Control **25**, 755 (1980).

[21] A. Scheinker and M. Krstic, Extremum seeking with bounded update rates, Systems and Control Letters **63**, 25 (2014).

[22] A. Scheinker, X. Pang, and L. Rybarcyk, Model-independent particle accelerator tuning, Phys. Rev. ST Accel. Beams **16**, 102803 (2013).

[23] A. Scheinker, S. Baily, D. Young, J. Kolski, and M. Prokop, In-hardware demonstration of model-independent adaptive tuning of noisy systems with arbitrary phase drift, Nucl. Instrum. Methods Phys. Res., Sect. A **756**, 30 (2014).

[24] P. Ioannou and J. Sun, *Robust Adaptive Control* (Prentice-Hall, Upper Saddle River, NJ, 1996).