Nuclear Inst. and Methods in Physics Research, A 1031 (2022) 166596

Contents lists available at ScienceDirect

Nuclear Inst. and Methods in Physics Research, A

journal homepage: www.elsevier.com/locate/nima

Design and verification of adaptive control to generate flat-top radio frequency power output of pulse compressor for SXFEL

Yiming Xu [a,b,c], Junqiang Zhang [c,*], Chengcheng Xiao [c], Wencheng Fang [b,c,*], Zhentang Zhao [c]

[a] Shanghai Institute of Applied Physics, Chinese Academy of Sciences, Shanghai 201800, China
[b] University of Chinese Academy of Sciences, Beijing 100049, China
[c] Shanghai Synchrotron Radiation Facility, Shanghai Advanced Research Institute, Chinese Academy Sciences, Shanghai 201204, China

ARTICLE INFO

Keywords:
SXFEL
Adaptive control
LLRF
RF pulse compressor

ABSTRACT

The adaptive control-based low-level radio frequency (LLRF) algorithm was developed for the two-bunch operation of the Shanghai Soft-X-ray Free Electron Laser Facility (SXFEL), needing to generate flat-top radio frequency (RF) power pulses at the output of an RF pulse compressor. The adaptive algorithm optimized for the compressor system can achieve a better convergence rate and domain. The algorithm has already modulated the flat-top-accelerating gradient in the SXFEL's RF cavity, and the energy of the electrons accelerated at the field's different longitudinal locations is within 0.8% (rms) of the mean value. This study presents the algorithm's theory, experimental setup, test results before and after the improvement, and the distortion of the measured waveform due to the transmission delay. Finally, the adaptive control algorithm-based LLRF system is installed at SXFEL and numerous results, including RF pulse stability and beam stability, are presented.

1. Introduction

The RF pulse compressor is critical equipment for compressing long low-power RF pulses to short high-power RF pulses [1], and the energy of such pulses is too high for klystrons to generate directly [2]. Presently, all pulse compressors installed and operating at SXFEL are in their original state [3], in which the output RF pulse is attenuated exponentially after the phase flipping. Because the energy of every bunch along the bunch chain is attenuated by the time variation within one RF pulse, which cannot radiate in the downstream undulator, the attenuated pulse cannot support multi-bunch or even more than two-bunch operations. Moreover, there is a large peak in the output RF pulse, which causes the accelerating gradient distribution in the accelerating structure to be nonuniform and significantly limits the maximum energy gain of each accelerating structure. In the upgrading plan, SXFEL will totally build six beamlines that can operate at 50 Hz repetition rate while being simultaneously driven by a single Linear Accelerator (LINAC) [4]. The LINAC's repetition rate is set to 50 Hz, which means that the LINAC must operate in multi-bunch modes to transport uniform energy along the bunch chain. To achieve uniform energy distribution in the bunch chain, the LLRF control algorithm must be upgraded to implement the compressor's flat-top output RF pulse, which also increases the LINAC efficiency in RF power and makes the bunch insensitive to timing jitter [3].

There are some methods [3,5,6], and [7] to generate flat-top RF power pulses at the output of an RF pulse compressor, but these algorithms must first acquire specific parameters, such as the pulse compressor's unload quality factor and its coupling coefficient, before calculating the LLRF output pulse based on the formulas in the above studies. However, because the RF system is time variable that is sensitive to the environment and manufacturing, it cannot be guaranteed that the real value is the same as the design specification. Considering all the factors above, a type of algorithm based on adaptive control with multivariable estimation is proposed and developed to provide a stable output for the SXFEL upgrading, which treats the RF system as a black box with only the input and output signals. This algorithm estimates the real-time mapping function between the inputs and outputs based on the real-time error between predicted and measured values. Furthermore, the adaptive algorithm could be applied directly in multiple compressor systems without configuring them separately, even if the parameters in these systems were significantly different, which is useful for large facilities. Given the benefits listed above, the algorithm is useful for free electron laser (FEL) [8] facilities such as the SwissFEL [9], SXFEL [10], and the SPring-8 Angstrom Compact free electron Laser (SACLA) [11].

The adaptive control algorithm-based LLRF system was developed, and it can both stabilize the amplitude within 0.08% (rms) and maintain the flat profile in 11 h. Moreover, it can stabilize the target automatically within 2 min when a large frequency offset occurs. In the SXFEL test, electrons with energies within 0.8% (rms) were observed accelerating at different longitudinal locations of the field modulated

* Corresponding authors.
E-mail addresses: zhangjq@sari.ac.cn (J. Zhang), fangwc@sari.ac.cn (W. Fang).

https://doi.org/10.1016/j.nima.2022.166596
Received 12 December 2021; Received in revised form 28 February 2022; Accepted 28 February 2022
Available online 16 March 2022

by the algorithm. The automatic energy switch achieves within 15 s. Furthermore, the waveform distortion was discovered during the experiment because the transmission delay is not integral multiples of the sampling period, and the corresponding solution was proposed. The principle of adaptive control is introduced in this study, followed by the experimental setup and results. The results include function implementation and continuous long-term test. The adaptive algorithm was improved in three directions and tested again. Section five presents the online results with beam test at SXFEL. Finally, at the end of the paper, the conclusion is drawn.

2. Adaptive control with multivariable estimation

The adaptive control can predict the dynamic behavior of the controlled object as well as system error and its variation in real time by measuring the system's input and output, and then adjusting the controller's parameters to maintain optimal system control and meet the predetermined requirements [12]. When the system's parameters change or are difficult to obtain, it is time for adaptive control to shine. Multivariable estimation could aid in determining the relationship between multiple independent inputs and outputs.

2.1. Algorithm flow

The main idea of the control algorithm is as follows.

(i) Initialize the relevant parameters.
(ii) Record the real output ($U(l)$ with dimensions of $n \times 1$) of LLRF system and measure the output ($Y(l)$ with dimensions of $m \times 1$) of the compressor.
(iii) Calculate the expected output ($\hat{Y}(l)$ with dimensions of $m \times 1$) of the compressor by the estimative mapping ($\hat{B}$ with dimensions of $m \times n$), $\hat{Y}(l) = \hat{B}U(l)$.
(iv) Modify the $\hat{B}$ according to the real-time error between the estimation $\hat{Y}(l)$ and the sample $Y(l)$.
(v) Solve equation $Y_r = \hat{B}U(l+1)$ to calculate the $U(l+1)$ that will be output by LLRF in the next pulse, where Y_r (a matrix with dimensions of $m \times 1$) is the set value of the expected compressor output.
(vi) Output $U(l+1)$ to the system, then back to step (ii).

2.2. Multivariable parameters estimation

This section will explain how to modify the mapping function in step (iv) and give the relevant formulas and theories.

There is an m-out and n-in linear system, $Y = BU$, and l groups of samples, $(Y_i(k), U_j(k))$, where $l > n \geq m$. The parameter matrix B can be estimated by these samples following [13]:

$$\hat{B} = YU^T \left(UU^T\right)^{-1}, \tag{1}$$

where U, Y, and $\hat{B}$ have dimensions of $n \times l$, $m \times l$, and $m \times n$, respectively. It can be proven that Eq. (1) is the least square estimation and the maximum likelihood estimation about B.

Nevertheless, Eq. (1) is impractical to update $\hat{B}$ in real time, which consumes a lot of memory and computational resource as the samples increase. In addition, the $\hat{B}$ changes more and more slowly as l, the number of samples, increases, which is the data saturation. Fortunately, the recursive least square method which updates $\hat{B}$ with forgetting factor λ ($0 < \lambda < 1$) can solve the above two problems by calculating Eq. (2), Eq. (3), and Eq. (4) [14,15] successively when a new sample is obtained.

$$K(l+1) = \frac{P(l)U(l+1)}{\lambda + U^T(l+1)P(l)U(l+1)} \tag{2}$$

$$P(l+1) = \frac{1}{\lambda}\left[I - K(l+1)U^T(l+1)\right]P(l) \tag{3}$$

$$\hat{B}(l+1) = \hat{B}(l) + \left[Y(l+1) - \hat{B}(l)U(l+1)\right]K^T(l+1) \tag{4}$$

where the $K(l)$ and $P(l)$ with dimensions of $n \times 1$ and $n \times n$ respectively are the intermediate variables to update $\hat{B}$ and I is the identity matrix of $n \times n$. Moreover, after introducing λ, the latest sample has a larger weight factor in the estimation of $\hat{B}$ than the previous samples. In the iterative calculation, only the real-time $\hat{B}(l+1)$ and $P(l+1)$ need to be saved until the next turn and the dimensions of all variables are constant.

2.3. Optimum control function

In step (v), the control goal is to minimize the following cost function:

$$J = E\left\{\left[Y(k) - Y_r(k)\right]^T\left[Y(k) - Y_r(k)\right]\right\},$$

where $Y(k) = BU(k) + \Xi(k)$ and $\Xi(k)$ is the white noise vector. The white noise is independent to any variable, so

$$J = E\left\{\left[BU(k) - Y_r(k)\right]^T\left[BU(k) - Y_r(k)\right]\right\} + E[\Xi^T(k)\Xi(k)].$$

If the solution of the linear system of equations, Eq. (5), exists, the solution is the optimum control that minimizes J.

$$Y_r = BU \tag{5}$$

3. The low-power test setup comprised pulse compressor and solid-state amplifier

A low-power RF experimental setup was built in reference to the framework of the SXFEL's RF unit [16]. However, there is not klystron in the setup, and the only thermostatic device is a water tank, ensuring the compressor temperature stable. The electronic devices and cable are explored in the atmosphere directly, so the setup is easily influenced by environment. Fortunately, the setup is only for the verification and comparison of the algorithm, and its instability is helpful to show the algorithm property.

3.1. Layout of the setup and software

Fig. 1 shows a schematic of the experimental setup as well as a picture of it. A vector signal generator (Rohde & Schwarz SMW200A) generates an RF signal source (5712 MHz). First, the signal is transmitted to the local oscillator and clock generator (LOCG) [17], which generates the clock (CLK), local oscillator (LO), and reference (REF) signals for the MTCA card [18]. The LOCG is a device that can keep the frequency of output to the frequency of input ratio constant. In other words, the frequency of the LOCG output varies proportionally to the frequency of the LOCG input, ensuring that the analog-to-digital convertor (ADC) rate is four times as frequent as the intermediate frequency (IF) signals accurately for in-phase quadrature (IQ) demodulation.

DWC8VM1HF [19] is an analog printed circuit board (PCB) that can convert signal frequencies. By mixing the RF output of the compressor with the LO signals (5685.5 MHz), the RF output is down-converted to IF signals (26.4 MHz). By mixing the 16-bit digital-to-analog convertor (DAC) output with the REF signals (5712 MHz), the vector modulator (VM) up-converts the IQ signals to the RF signals. The digital PCB SIS8300L2 [19] is linked to the analog board via the internal bus. 16-bit ADC samples the IF signals with the CLK signal (105.7 MHz). As the compressor's input, the solid-state amplifier (SSA) amplifies the output of the MTCA to hundreds of watts.

Experimental physics and industrial control system (EPICS) is used to read all samples, to compute intermediate variables, and to write the feedforward table in LLRF. The LLRF communicates with another device via the local area network, allowing the external terminal to read and write process variables (PV) in EPICS. The amplitude control algorithm was created on the Python [20] platform, which was installed on the external terminal. It reads the required PVs from EPICS, calculates the control values, and then returns them. This process is repeated at

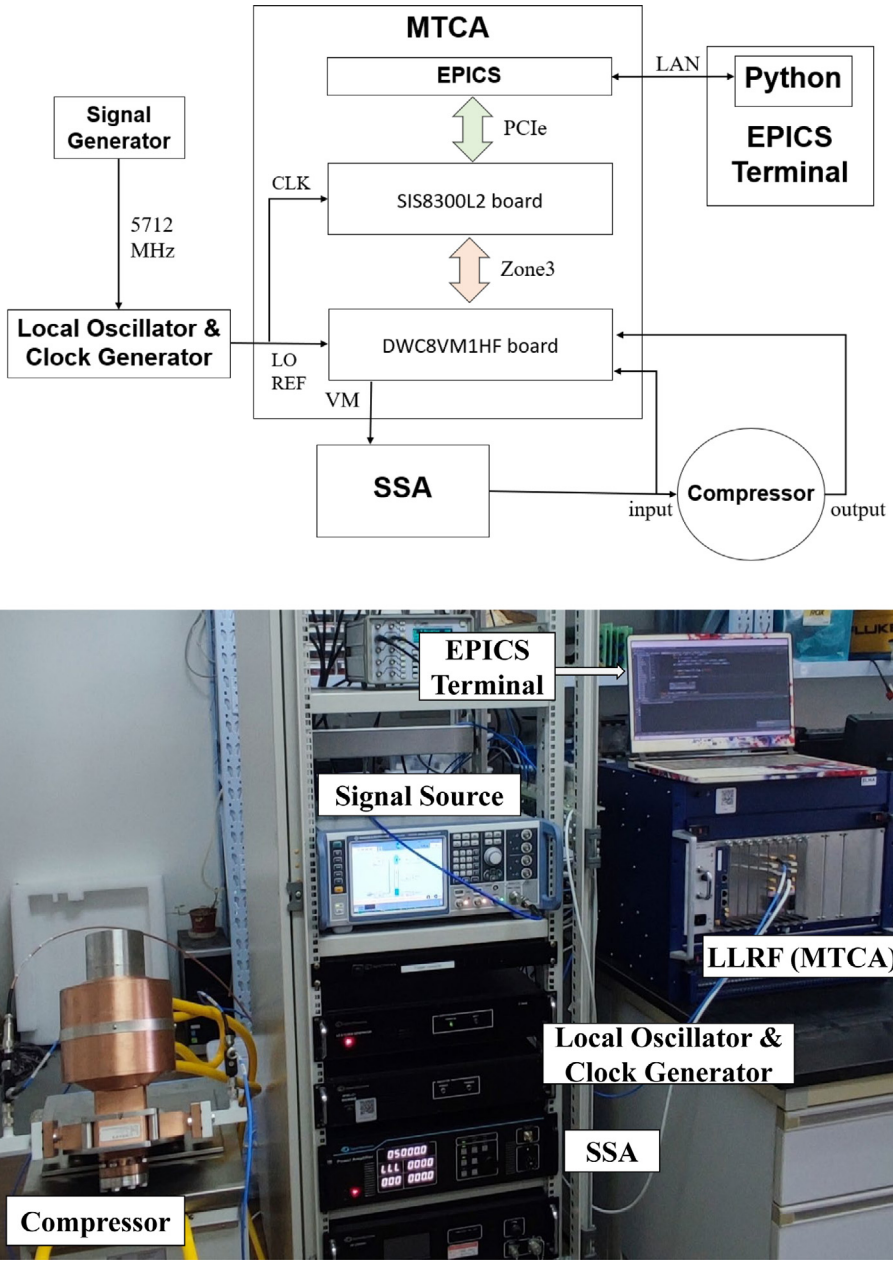

Fig. 1. Schematic (up) and a picture (down) of the experimental setup. (Key devices are marked in the corresponding positions of the diagram).

a rate of 1 Hz. The control is achieved by changing the feedforward table in real time. The compressor is a C-band compact spherical RF pulse compressor [21] with parameters that differ significantly from the theoretical values. The compressor is connected to a thermostatic water tank.

3.2. Control method and mathematical model

When the RF frequency matches the resonant frequency of the compressor, the output phase is flat spontaneously, which does not require special algorithm. As for the setup, the RF frequency from the signal generator is adjustable. In SXFEL, the resonant frequency can be adjusted and maintained by the thermostatic water system. Furthermore, an independent feedback loop embedded in LLRF employs the PID [22] algorithm to stabilize the output phase against the slow drift. The control value of phase is calculated according to the error between the mean value of flat-top phase and the set phase. The output phase of LLRF is constant within a pulse. In the original state, the output amplitude is not flat, which needed to be modulated. Because the formula-based amplitude modulation [5] depends on the compressor's parameters that are sensitive to the environment, a method, not requiring the specific parameters, is wanted. Therefore, the adaptive control was employed to flatten the output amplitude alone. Fig. 2 shows the schematic of the feedback loop of amplitude and phase.

The total period of the input pulse is 2.44 μs, with a charging period of 2 μs and a discharging period of 0.44 μs (shown in Fig. 3). When the input phase flips by 180°, the charging period is over and the discharging period begins. The target is flattening the output during the discharging period. In the discharging period, controlling the input amplitude waveform can modulate the output amplitude waveform. To sample the IF signals, the 105.7 MHz ADC employs IQ sampling, yielding one sample every 40 ns. It should be accepted because the first output sample near the positive edge cannot reach the set value. As a result, the output is quantized to a 10-sample sequence. The

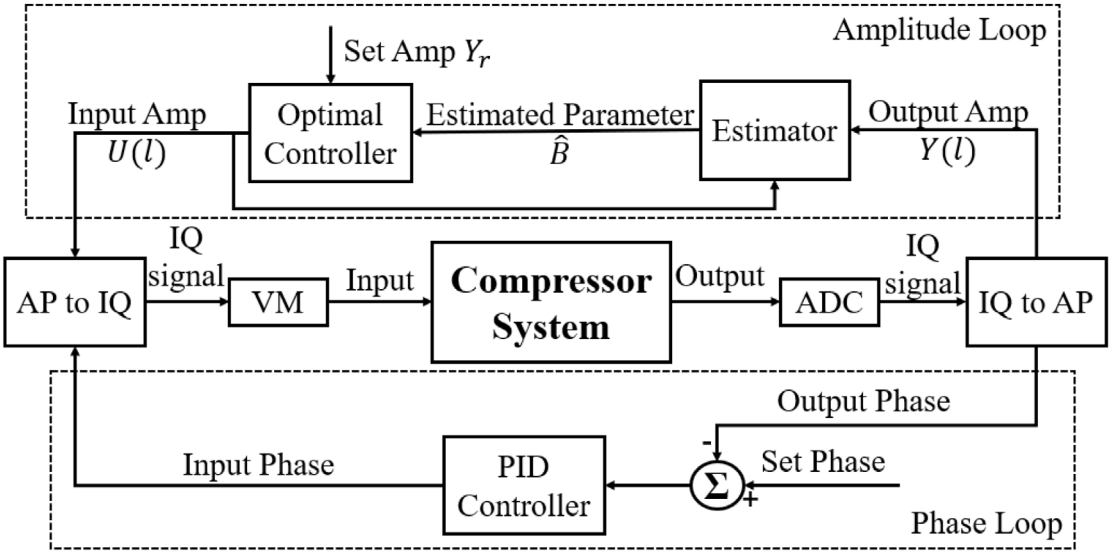

Fig. 2. Schematic of feedback loop of amplitude (up) and phase (down).

compressor's input during the discharging period is also described by a 11-sample sequence with 40 ns intervals. Adjusting the 11 phase-flipped input points could effectively control the above 10 output points. The pulse interval is 1 s, which is much longer than the pulse length, so adjacent pulses do not interfere with each other. Furthermore, if the system is stable, the inputs and outputs always change in a narrow range; therefore, their relationship is approximated as a linear mapping. Finally, the system has 11 inputs, 10 outputs, is zero-order, and is linear. The following function can represent the system's mathematical model:

$$Y(k) = z^{-1}BU(k) + b + \Xi(k), \tag{6}$$

where $\Xi(k)$ is the white noise vector and b is the offset vector. The 10×11 matrix B and the 10×1 vector b are the parameters needed to be estimated by the adaptive control.

3.3. Performance limitation of DAC rate

As is well known, discrete digital systems only behave as continuous systems when the time interval is short enough and the bits of quantization are sufficient. The DAC in this LLRF keep their outputs constant within 40 ns, indicating that the VM output is a staircase function. If the input is constant, the output after 180° phase flipping will decrease, according to Eq. (7) [5].

$$T_c \frac{dE_{out}}{dt} + E_{out} = -T_c \frac{dE_{in}}{dt} + \frac{\beta - 1}{\beta + 1} E_{in}, \tag{7}$$

where T_c is the filling time of compressor, β is the coupling coefficient of compressor, E_{in} is the amplitude of the input RF field, and E_{out} is the amplitude of the output RF field.

In this configuration, the compressor has a T_c of approximately 0.9 μs and a β of approximately 4.4. As shown in Eq. (7), the compressor output will decrease by 6% in 40-ns intervals. However, because the bandwidth of a high-power klystron and an RF pulse compressor are comparable to the present DAC and ADC rates, the oscilloscope-measured decline is only 2% (the red line in Fig. 9 shows the result). Consequently, increasing the DAC rate is an effective way to increase the upper limit of the output flatness.

4. Verification and improvement of algorithm

At the above setup, the adaptive algorithm was tested and verified. The algorithm only changes the input amplitude, so this section focuses on the stability and flatness of the amplitude. The amplitude unit of all variables measured by LLRF is the ADC counts. The flatness of the flat-top output is expressed by the stand deviation of the samples during the discharging period.

4.1. Flat-top pulse generated by adaptive control with multivariable estimation

As for the RF system, an 11-in and 10-out system, the ith output will not be influenced by the jth input if i is not larger than j. therefore, B in Eq. (6) should be a lower triangular matrix, so the shape of the estimation $\hat{B}$ is hoped the same as B. If $\hat{B}$ is updated as a whole, its shape will be changed by the iterative calculation. Thus, the 11-in and 10-out system can be regarded as 10 i-in and single-out subsystems ($i = 2, 3, \ldots,$ and 11) with $(i+1)$ parameters. The jth subsystem is only relative to the jth parameter of b and the jth row of B ($j = 1, 2, \ldots,$ and 10). They are calculated independently according to Eq. (2), (3), and (4), and then these updated parameters are put back where they were.

In Fig. 3(a) and (b), the red line represents the compressor's original output amplitude and phase, and the blue line represents the flat-top output generated by the adaptive algorithm, which was the first time the flat-top output was acquired. The overshot in the flat-top phase is the waveform distortion because IQ demodulation is not good at measuring the signal's rapid change. The mean value of the flat top is 23010 with −26.8°. The amplitude flatness is 0.05% (rms), which was primarily influenced by the energy of white noise. The phase flatness is 0.5° (rms), which will be better if the RF frequency was more precisely adjusted.

4.2. Improvement of convergence rate

To observe the long-term behavior of the algorithm, a 4.5-h continuous experiment was done, shown by Fig. 4(a) and (b). The default value is 26500. They represent the mean value and standard deviation of the 10 outputs. The mean value varied within 0.11% of the set value and the standard deviation varied within 0.11% (rms). Initially, the output converged to the set value in 50 s, resulting in 50 iterations. However, there were numerous spikes in Fig. 4. The amplitude of the spikes was approximately 1000 and the width was about 5 min. The reason why these spikes occurred is explained below.

When the setup changes a little due to the environment, the estimation, $\hat{B}$, should be changed to follow the system change. Nevertheless, the different parameters in $\hat{B}$ have different importance for the system,

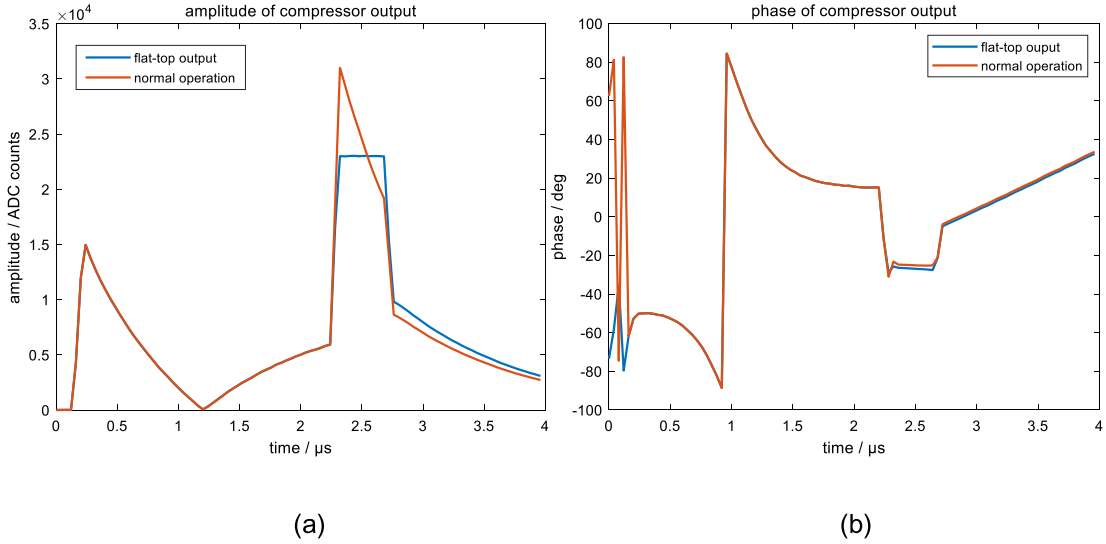

Fig. 3. Flat-top output of compressor by adaptive control.

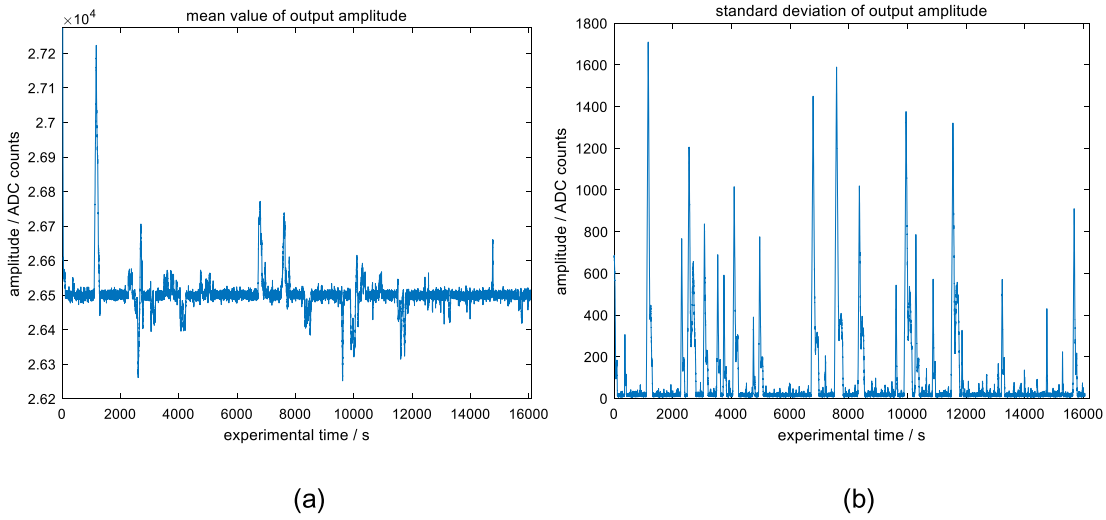

Fig. 4. Continuous test of flat-top amplitude.

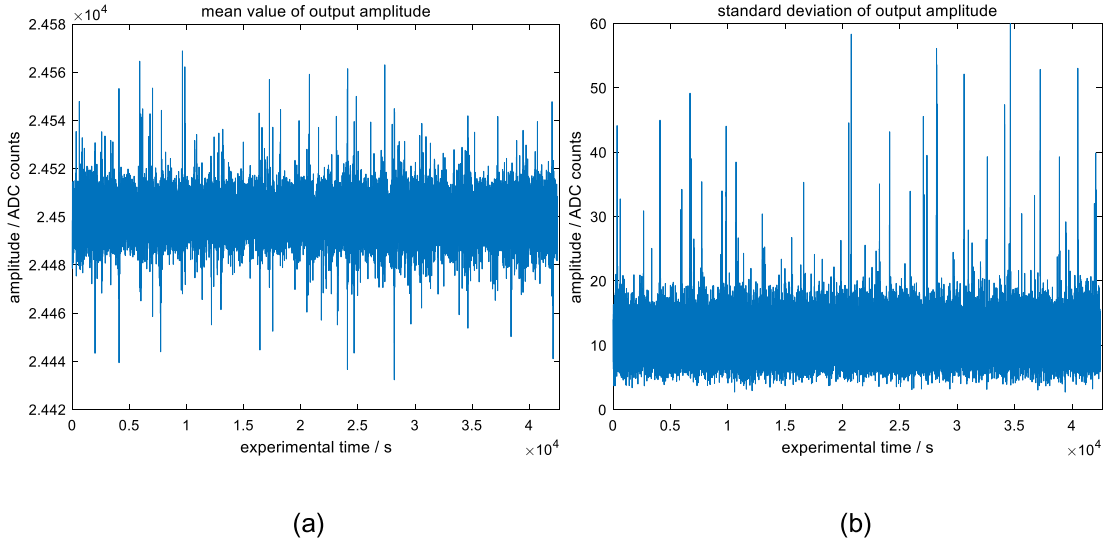

Fig. 5. Continuous test after improvements.

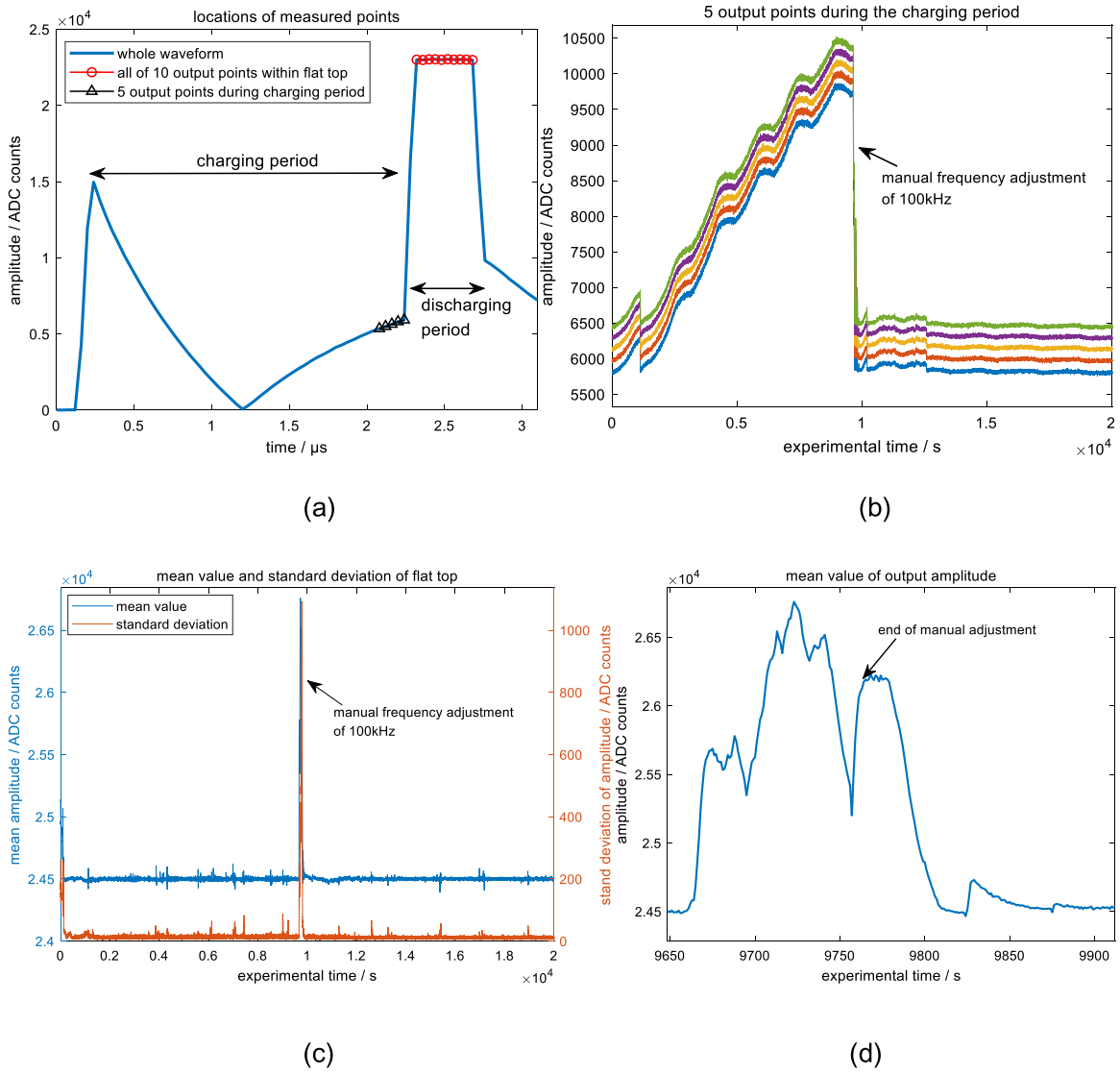

Fig. 6. Continuous measurement about output points that are marked in (a). (b) is the long-term test of 5 output points without control that are marked in (a) by the black triangles. They varied following the system status. (c) is the long-term test of 10 output points that are marked in (a) by the red dots. They have always been controlled by the algorithm. (d) is the enlarged figure of the blue line in (c) at the highest peak.

but the algorithm does not know that. When the algorithm needs to re-estimate too many parameters, it may make some more significant parameters further away from their actual values at the beginning. Therefore, the control value calculated by $\hat{B}$ makes the system output further away from the set value at beginning, which amplifies the setup change. So, the spikes appear and have large amplitude.

Moreover, the greater the number of parameters in a system, the more samples are required to estimate these parameters. For example, the tenth subsystem of the 10 i-in and single-out subsystems has 12 parameters. These parameters may have changed significantly before an enough number of samples were obtained. In other words, especially for multi-parameter and time-variable systems, the number of samples is insufficient to keep the estimation error within an acceptable range before slow drift becomes apparent. This makes the spikes have large width.

According to the above experiment, the adaptive algorithm could correct spikes finally every time but the algorithm's convergence time needs to be reduced. As a result, the adaptive control algorithm has potential and could be improved in the three areas listed below.

To begin with, B is a lower triangular matrix; however, the closer the matrix elements are to the bottom left, the lower their values are compared to those on the leading diagonal. As a result, the lower-left elements have no bearing on the calculation of the control values. In addition, they slow the rate of estimation convergence. As a result, if $\hat{B}$ is defined as a band matrix [23] with three elements in each row except the first, the convergence rate will be accelerated without compromising the estimation accuracy. According to the previous experimental result, the second output point with three parameters is the most stable of the 10 points so each row of $\hat{B}$ keeps three elements. Furthermore, because the improved algorithm reduces algorithmic time complexity from $O(n^2)$ to $O(n)$, it is simple to spread to systems with higher sampling rates.

Second, the forgetting point, λ, was converted from a constant to a variable.

$$\lambda(k) = a + \frac{(1-a)\,c}{c + \left|Y_r - \overline{Y}(k)\right|},$$

where $a \in (0,1)$, c is a positive real number and $\overline{Y}(k)$ is the mean value of output. This study chose $a = 0.95$ and $c = 200$. The larger the estimative error, the smaller λ is, and vice versa. At a smaller λ, reducing the weight of the previous samples would speed up the algorithm's tracking of parameter variation, while larger λ would strengthen the algorithm's resistance to white noise.

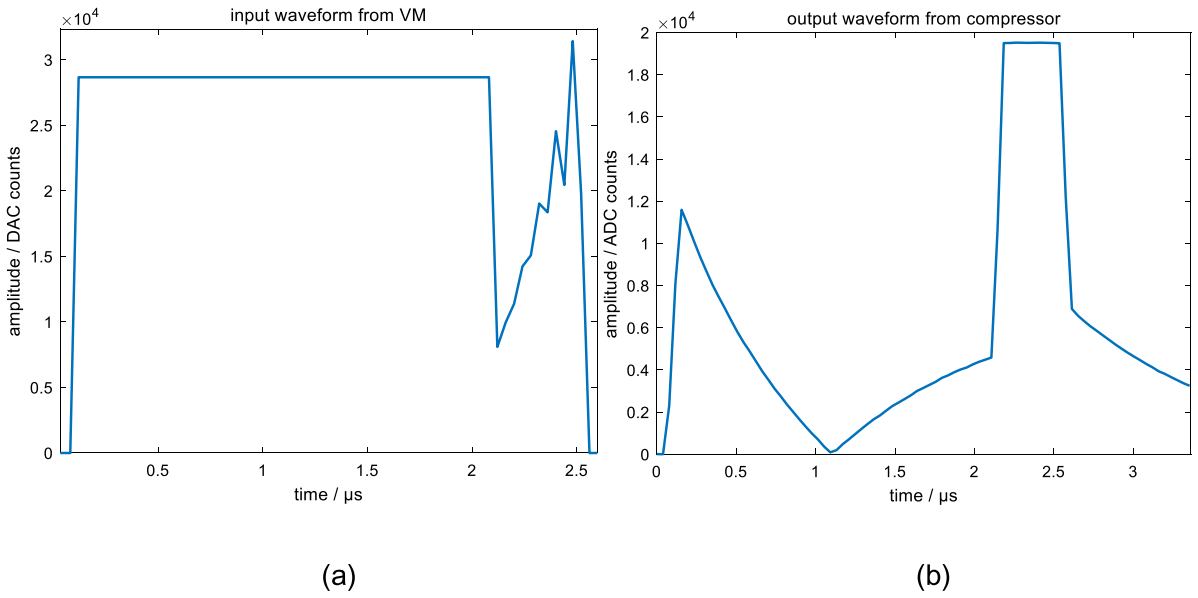

Fig. 7. The unusual input and output waveform measured by LLRF.

Third, for this RF system, B is a 10×11 matrix, indicating that there are 10 equations and 11 unknowns, implying that the solution set of the linear system of equations (5) is a straight line in an 11-dimensional space. In the previous experiment, the control point on this straight line was chosen at random. If B were constant, all points on this straight line or in the solution set would be the optimal control values. However, the RF system is not linear in and of itself and it can only be considered a linear system only if it always performs in the neighborhood of the steady point. In other words, B is a function of the input points U. If the control point U chosen on this straight line is far from the last control point U_0, B would change dramatically but $\hat{B}$ would not, resulting in a large estimative error. Because the algorithm must recalculate B, spikes appear. Considering the preceding problem, the point in the solution set closest to U_0 should be chosen as the control point by Eq. (8), and the point is one and only one.

$$U = U_0 + B^T \left(BB^T\right)^{-1} \left(Y_r - b - BU_0\right) \tag{8}$$

Fig. 5 depicts the outcome of an 11-h experiment after the three improvements mentioned above. The set value is 24500. Obviously, there are fewer spikes than those in Fig. 4. The spike width is reduced from 5 min to 4 s, the amplitude of the spikes is reduced from 1000 to 70, the variation in the mean value is reduced to 0.08% of the set value, and the standard deviation is reduced from 0.11% to 0.08% (rms). Because the setup is more easily influenced by the environment, the spikes still appear frequently. If experiment on a steady facility, the better result will be got.

Later on, to observe how the algorithm deals with the extreme change of the system status, the thermostatic water tank was removed to make the system status vary constantly following the atmosphere temperature. Fig. 6 depicts the outcome of a 5.5-h experiment. The default value is 24500.

When the compressor status varied following the temperature, five black triangle points out of the feedback control area are also varied following the status. Fig. 6(b) shows their variation. The signal generator was responsible for the ripple effect. The uptrend in amplitude was caused by the variable system status. The upheaval occurred because the RF frequency was manually readjusted to match the resonance frequency at the time. Fig. 6(c) shows the mean amplitude and standard deviation of all output points in the discharging period (marked in Fig. 6(a) by the red dots) respectively, throughout the 5.5-h experiment. In stark contrast to Fig. 6(b), there is no ripple or slow drift in (c). Furthermore, the output turn back to the set value automatically after

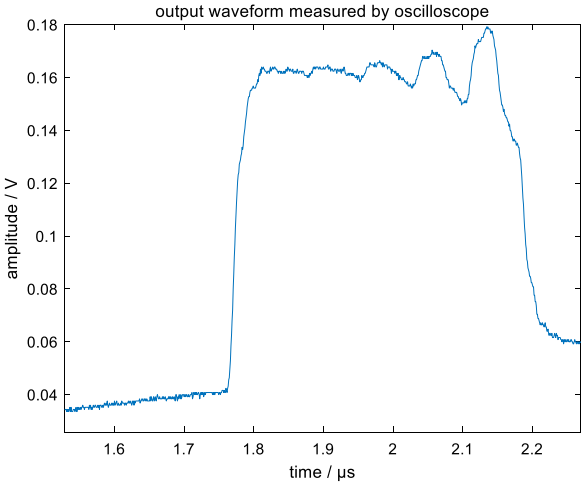

Fig. 8. Output waveform measured by the oscilloscope.

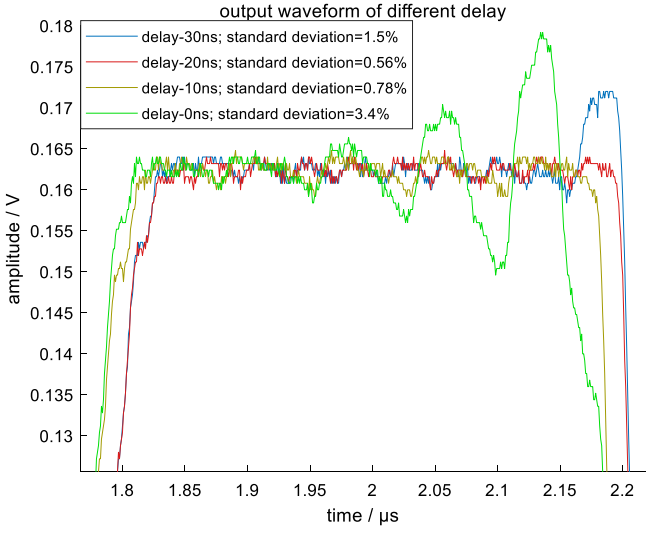

Fig. 9. Flat-topped output with different transmission delay.

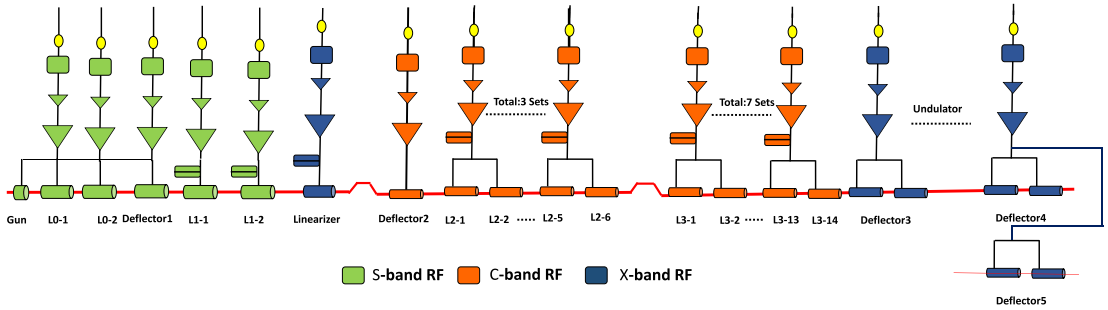

Fig. 10. Layout of RF units at SXFEL.

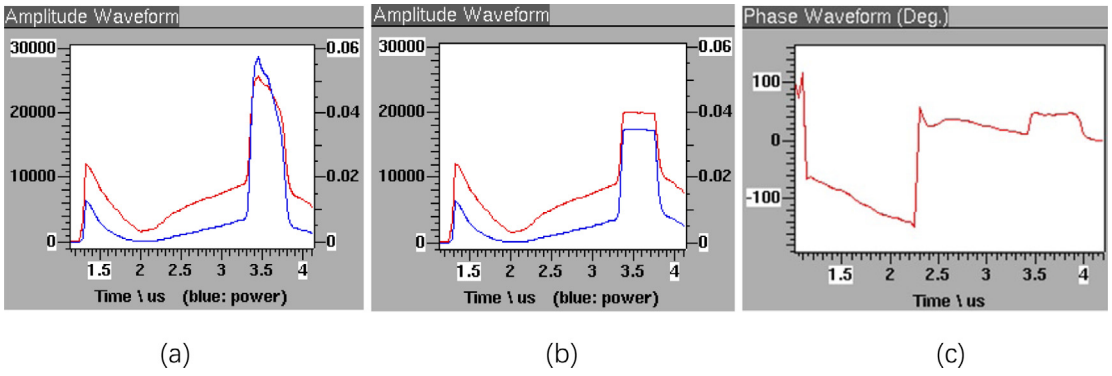

Fig. 11. Function implementation. (The (a) is the original output, (b) is the flat-top output, and (c) is the output phase. In (a) and (b), the red lines represent the output voltage, and the blue lines represent the output power.)

the manual frequency adjustment. Fig. 6(d) is the variation of the mean amplitude of flat top at the human intervention. The algorithm spent 50 s re-estimating the parameters and restoring the output to the set value within 50 iterations. This result demonstrates that the adaptive control has a fast convergence rate and a large convergence domain. Up to now, the verification of the algorithm at the setup is finished, and the acceptable result has been acquired. The test at SXFEL is the next step.

4.3. Waveform distortion caused by transmission delay

An unusual result was discovered in some experiments. Even when the input waveform was saw shaped, the output was quite flat, as shown in Fig. 7.

At the same time, the output was also sampled by a 1.25-GHz oscilloscope (Tektronix DPO 4054B) with an additional detector. Fig. 8 shows this result, which is quite different from the measurement of LLRF.

It takes some time that the VM output is transmitted to the output port of compressor via the coaxial cables, SSA, and compressor. Then, it is sampled as the compressor output. This is the transmission lag time. The same CLK signal activates both the ADC and the DAC in the LLRF. Suppose the delay is not an integral multiple of the sampling period of 40 ns. In that case, the DAC trigger edges that have experienced the transmission delay will be misaligned with the ADC trigger edges. The greater the misalignment, the greater the distortion, and thus the controller receives the wrong message. This is why the above-mentioned unusual phenomenon appears.

For IQ demodulation, a sample is obtained every 4 CLK periods of 10 ns, allowing LLRF to adjust the transmission delay with an accuracy of 10 ns, resulting in a misalignment of no more than 5 ns. The delay was reduced by 30 ns, 20 ns, 10 ns, and 0 ns in the following experiments. The adaptive algorithm then began to compute the control values, and LLRF acquired the flat-top output at different delays. Finally, the oscilloscope was used to sample the outputs that were considered flat by LLRF. The outcome is depicted in Fig. 9. The output standard deviations are labeled in Fig. 9. The least misalignment was obtained when the delay was reduced by 20 ns (the red line).

Because the LLRF ADC rate is insufficient, all of the standard deviations measured by oscilloscope were greater than 0.08%. The red line decreased by 2% per 40 ns. The ripple is inevitable when the digital systems are used to flatten the compressor output (explained in Section 3.3). The IQ demodulation averaged the output over a sampling period equal to the above decline time, so LLRF cannot detect the large ripple in the real waveform. By adjusting the length of the cable, the misalignment can be reduced to less than 5 ns and the changed length will not be greater than 1.5 m. Besides that, the transmission delay can also be adjusted more accurately by faster ADC and DAC.

Usually, the transmission delay of the system does not change obviously over time. Therefore, the oscilloscope is needed to judge the misalignment only at the initial step. In the operating status, the adjacent bunches' energies and the output waveform of LLRF can also imply the accuracy of the flatness.

5. Experimental results to verify the two-bunch operation at SXFEL

The SXFEL facility has 10 C-band RF units. Fig. 10 shows the layout of all RF units at SXFEL. The adaptive algorithm was used to flatten the output amplitude of the last C-band RF unit. This unit has an additional klystron, and it uses a SLED-type compressor. The electronic devices are installed in a thermostatic cabinet, and the cables and RF cavities are heat-preserved by the thermostatic water system that can adjust the resonant frequency of compressor to match the RF frequency. The unit is much steadier than the previous experimental setup. Before experiment, the interface and transmission delay were adjusted. The result is derived from both the RF compressor's pulse output and the energy of the electron beam.

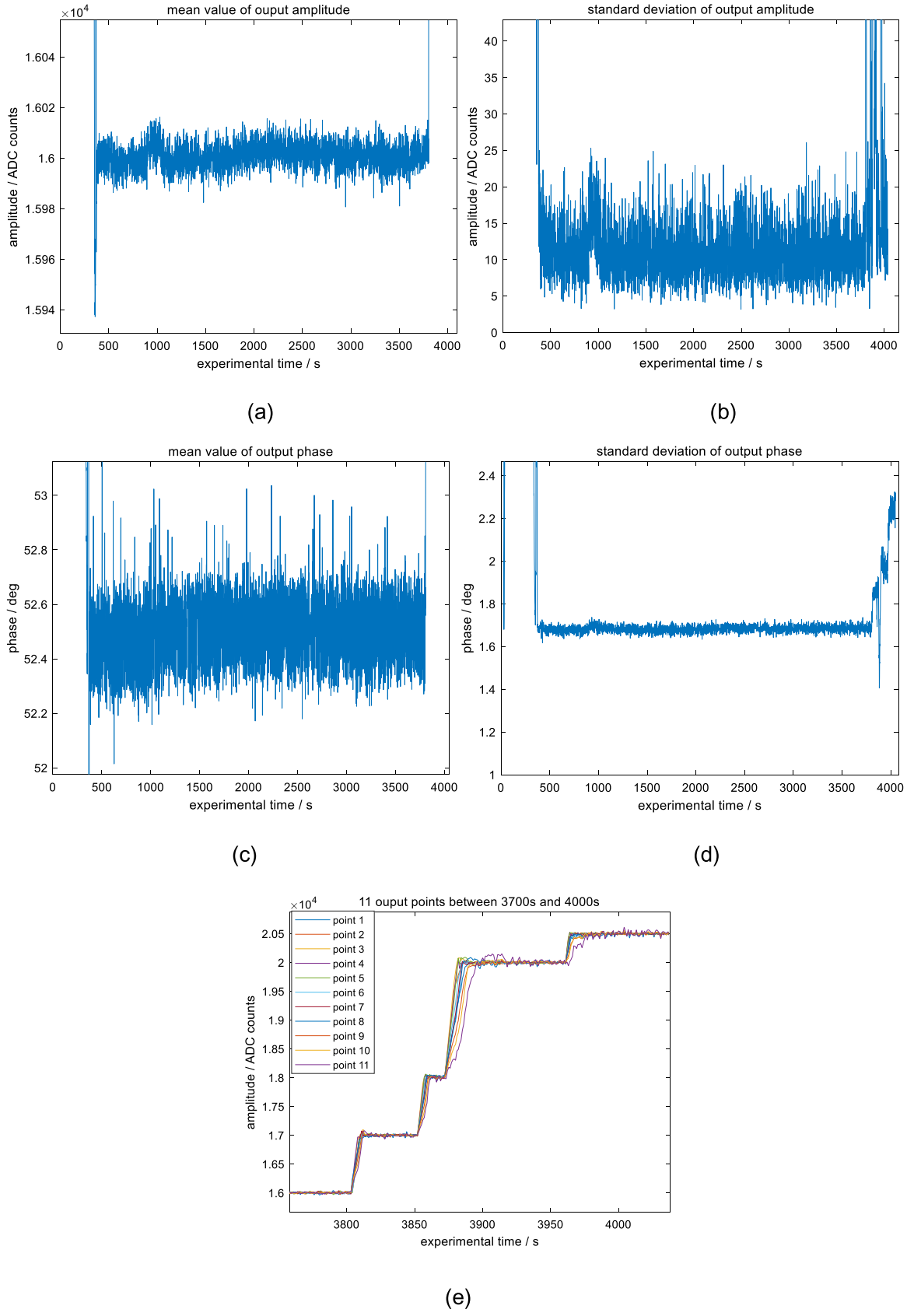

Fig. 12. Long-time stability and dynamic property at SXFEL.

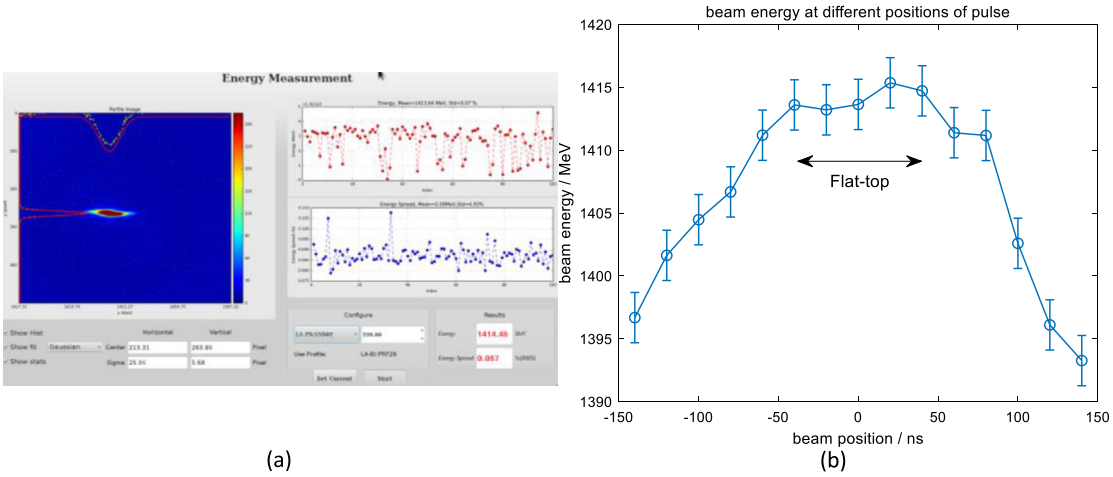

Fig. 13. The stability of beam energy and the variation of beam energy with the trigger change. (The flat top contains five points).

5.1. Function implementation, long-term stability, and switch time of RF power

After several minutes of estimation, the amplitude of compressor output in this unit was transformed from Fig. 11(a) to Fig. 11(b). Fig. 11(c) is the output phase after the modulation. Because the klystron outputs at SXFEL have a sharp positive edge, the flat top have 11 output points. Consequently, the algorithm is now available for SXFEL.

Next, the algorithm's long-term stability was tested. Fig. 12(a) and (b) show the mean and standard deviation of flat-top amplitude, respectively, in 1 h, with a set value of 16000. The set value was manually changed at the end of the test to observe the dynamic behavior of the algorithm, so the upheaval occurred at the end. Fig. 12(c) and (d) show the mean and standard deviation of flat-top phase, respectively. Fig. 12(e) depicts the variation of 11 outputs during the human intervention. Because the SXFEL environment is relatively stable, the spikes in the table experiment vanish in Fig. 12(a) and (b). The mean amplitude variation is 0.09% of the set value, and the standard deviation is 0.13% (rms). The mean phase varies in 52.45±0.25° and the standard deviation is 1.7° (rms).

The algorithm spent the first 7 min estimating the mapping between the LLRF output and the compressor output. The mapping was saved and would be loaded directly the next time, flattening the output in 1 min. The switch time in Fig. 12(e) is approximately 15 s. The estimation rate is adequate. If the resonant frequency of compressor is adjusted to match the RF frequency more precisely, the better flatness of phase will be acquired.

5.2. Flatness test based on beam energy

SXFEL cannot currently operate in multi-bunch mode. The beam energy was measured when the triggers of the RF pulses were changed while the triggers of the bunches were fixed to ensure whether the beams at different positions experienced the same accelerating gradient. The relationship between time and accelerating gradient could be discovered in this manner. Furthermore, the step of change is 20 ns. The compressor's output length is 440 ns, and the accelerating structure's filling time is 360 ns. As a result, the accelerating interval should have four or five bunches. Fig. 13(a) depicts bunch stability, with 90% of beam energy falling within 2 MeV of the mean value. The error bars in Fig. 13(b) represent a 90% range of beam energy. The experimental result is shown in Fig. 13(b).

The accelerating interval has five points. Their average energy is 1414.1 MeV. The energy gain of a C-band RF unit is about 100 MeV, so the flatness of bunches in flat top is 0.8% (rms). Peak-to-peak energy is 2.2 MeV. The focusing magnets were not adjusted when the RF pulse triggers were changed, so the facility did not process at the peak efficiency. In addition, the compressors of other RF units were working at the normal operation, making the bunches sensitive to the time jitter. Therefore, the bunch stability is not good enough. If the flat top is lengthened or the transmission delay is adjusted more precisely or all of compressors are flattened, the flatness will be improved.

6. Conclusion and outlook

The adaptive control-based algorithm has been designed, implemented, and validated in SXFEL. The results demonstrate that the algorithm can generate the expected flat-top RF power pulse of the compressor as well as stabilize long-term amplitude and flatness against environmental variations. It has a larger convergence domain and a faster convergence rate. Furthermore, the adaptive control algorithm is universal in that it is unaffected by specific system parameters. Finally, the energies of bunches in the accelerating interval are within the range of 2.2 MeV. Apart from that, it was discovered that the real output may not be as flat as the one obtained by LLRF and that changing the transmission delay can remit this problem. This study focuses on the algorithm's verification. The further research and mature proposal are in the schedule for the long-term operation, embedding in all RF units at SXFEL. The faster ADC and DAC will be applied in the LLRF control systems. Besides flattening the compressor output, adaptive control based on estimation and correction could be used in other accelerator applications, such as ultra-high dose rate (FLASH) proton therapy [24].

CRediT authorship contribution statement

Yiming Xu: Methodology, Writing – original draft, Data curation, Formal analysis. **Junqiang Zhang:** Software, Validation. **Chengcheng Xiao:** Validation, Funding acquisition. **Wencheng Fang:** Funding acquisition, Supervision, Writing – review & editing. **Zhentang Zhao:** Supervision, Methodology, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This document is the results of the research project funded by the Alliance of International Science Organizations, China (ANSO-CR-KP-2020-16) and the National Natural Science Foundation of China (No. 12105345).

References

[1] Zongbin Li, et al., RF design of a C-band compact spherical RF pulse compressor for SXFEL, Nucl. Instrum. Methods A 863 (2017) 7–14.
[2] Cheng-Cheng Xiao, et al., Design and preliminary test of the LLRF in C band high-gradient test facility for SXFEL, Nucl. Sci. Tech. 31 (2020) 100, http://dx.doi.org/10.1007/s41365-020-00806-6.
[3] Amin Rezaeizadeh, et al., Pulse Compressor Phase and Amplitude Modulation Based on Iterative Learning Control, JACoW-IPAC2015-MOPTY060, 1076-1078.
[4] Zhao Zhentang, et al., SXFEL: A Soft X-ray free electron laser in China, Synchrotron. Radiat. News 30 (2017) 29–33.
[5] Wang Chaopeng, et al., Design and study of C-band RF pulse compressor for sxfel linac, Nucl. Sci. Tech. 25 (2) (2014).
[6] R. Bossart, P. Brown, J. Mourier, I.V. Syratchev, L. Tanner, High-power microwave pulse compression of klystrons by phase-modulation of high-Q storage cavities, CLIC-NOTE-592, 2004.
[7] S.H. Shaker, et al., Phase Modulator Programming to Get Flat Pulses with Desired Length and Power from the Ctf3 Pulse Compressors, JACoW-IPAC2010-TUPEA043, 1425-1427.
[8] Zhang Junqiang, et al., A beam energy feedback for ultrafast electron diffraction facility, High Power Laser Part. Beams 32 (6) (2020).
[9] Eugenio Ferrari, et al., The ACHIP experimental chambers at the Paul Scherrer Institut, Nucl. Instrum. Methods A 907 (2018) 244–247.
[10] Wencheng Fang, et al., Design, fabrication and first beam tests of the C-band RF acceleration unit at SINAP, Nucl. Instrum. Methods A 823 (2016) 91–97.
[11] Xiaoxia Huang, et al., Design of an X-band accelerating structure using a newly developed structural optimization procedure, Nucl. Instrum. Methods A 854 (2017) 45–52.
[12] I.D. Landau, Commande Adaptative Aspects Pratiques Et Theoriques (Chinese Translation), Beijing Institute of Technology Press, Beijing, 1992.
[13] John A. Rice, Mathematical Statistics and Data Analysis, third ed., Thomson Brooks, US, 2007, pp. 255–311.
[14] Wikipedia: Recursive least squares filter. https://en.wikipedia.org/wiki/Recursive_least_squares_filter.
[15] Arvind. Yedla, A Tutorial on Recursive methods in Linear Least Squares Problems. http://pfister.ee.duke.edu/courses/ece586/ex_proj_2008.pdf.
[16] W. Fang, et al., RF System for SXFEL: C-band, X-band and S-band, JACoW-IPAC2018-THPMK064, 4446-4448.
[17] Chengcheng Xiao, et al., Design and test of a c band local oscillator and clock device in LLRF, Nucl. Tech. 44 (1) (2021).
[18] Tomasz Jezynski, et al., ATCA/μTCA for physics, Nucl. Instrum. Methods A 623 (2010) 510–512.
[19] Junqiang Zhang, et al., A precision LLRF control system for UED, Nucl. Instrum. Methods A 1021 (2021) 165587.
[20] https://www.python.org/.
[21] Z.B. Li, W.C. Fang, Q. Gu, et al., Design, fabrication, and cold-test results of C-band spherical RF pulse compressor prototype, Radiat. Detect. Technol. Methods 3 (2019) 21.
[22] Wikipedia: PID controller. https://en.wikipedia.org/wiki/PID_controller.
[23] Wikipedia: Band matrix. https://en.wikipedia.org/wiki/Band_matrix.
[24] W.C. Fang, X.X. Huang, J.H. Tan, et al., Proton linac-based therapy facility for ultra-high dose rate (FLASH) treatment, Nucl. Sci. Tech. 32 (2021) 34, http://dx.doi.org/10.1007/s41365-021-00872-4.