The submitted manuscript has been created by UChicago Argonne, LLC, Operator of Argonne National Laboratory ("Argonne"). Argonne, a U.S. Department of Energy Office of Science laboratory, is operated under Contract No. DE-AC02-06CH11357. The U.S. Government retains for itself, and others acting on its behalf, a paid-up nonexclusive, irrevocable worldwide license in said article to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government.

ANL/APS/LS-326
March 21, 2011

Dielectric Wakefield Accelerator to Drive the Future FEL Light Source

C. Jing[1,2], J. Power[1], and A. Zholents[3]

[1.] *High Energy Physics Division, ANL*
[2.] *Euclid Techlabs, LLC*
[3.] *Advanced Photon Source, ANL*

Abstract: *X-ray free-electron lasers (FELs) are expensive instruments and a large part of the cost of the entire facility is driven by the accelerator. Using a high-energy gain dielectric wake-field accelerator (DWA) instead of the conventional accelerator may provide a significant cost saving and reduction of the facility size. In this article, we investigate using a collinear dielectric wakefield accelerator to provide a high repetition rate, high current, high energy beam to drive a future FEL x-ray light source. As an initial case study, a ~100 MV/m loaded gradient, 850 GHz quartz dielectric based 2-stage, wakefield accelerator is proposed to generate a main electron beam of 8 GeV, 50 pC/bunch, ~1.2 kA of peak current, 10×10 kHz (10 beamlines) in just 100 meters with the fill factor and beam loading considered. This scheme provides 10 parallel main beams with one 100 kHz drive beam. A drive-to-main beam efficiency ~38.5% can be achieved with an advanced transformer ratio enhancement technique. rf power dissipation in the structure is only 5 W/cm^2 in the high repetition rate, high gradient operation mode, which is in the range of advanced water cooling capability. Details of study presented in the article include the overall layout, the transform ratio enhancement scheme used to increase the drive to main beam efficiency, main wakefield linac design, cooling of the structure, etc.*

I. Motivation

The free electron laser (FEL) is considered to be the main candidate for a short wavelength (UV to X-ray), short pulse (femto- to attoseconds) light source. Demands of the beam to drive such FELs have become more and more challenging, including high repetition rate (~MHz), high peak current (a few kA), and low emittance (sub-micron normalized emittance). More importantly, as a facility for other research, the light source has to be more economical and efficient. One approach to reduce the cost is to use a high

gradient beam acceleration system so that the electron energy can be boosted to the few GeV level (to obtain a high brightness beam) in a short distance. In the past decade, great progress has been made in the development of the FEL as the 4th generation light source. A number of facilities are in operation around the world and more have been planned [1-6]. During the same period of time, the field of high gradient acceleration, aimed at a future high energy collider, has also achieved impressive results [7-9]. Mutliple GV/m gradients have been demonstrated in various wakefield acceleration schemes. It is natural to think about using the state of art in high-gradient technologies as the backbone of the next generation FEL based light source to make the future FEL more compact and cost effective.

Unlike high energy physics machines, which run in pulsed mode, CW (i.e. continuous high repetition beam pulse) operation is preferred for the FEL light source. A real challenge for the accelerators in the future FEL application is to combine high gradient (>100 MV/m) with high repetition rate (~MHz). This combination rules out conventional klystron driven metallic accelerators since neither the normal nor the superconducting accelerator will be able to exceed a gradient of 100MV/m in the foreseeable future. High repetition operation also rules out the possibility of using the Two-Beam Accelerator (TBA) scheme because the main acceleration beamline of this scheme is essentially using a conventional technology (refer to CLIC [10]) which has yet to demonstrate to be able to maintain both high gradient and MHz repetition rates. The collinear wakefield acceleration scheme may work because it directly transfers energy from the drive beam's rf to the main beam. This rf pulse can be very short so high repetition rate may not trigger other problems, like rf breakdowns.

In this note, as a way of starting the discussion, we propose a straw man scheme using a collinear dielectric wakefield acceleration scheme to generate an 8 GeV, 50 pC/bunch, ~1.2 kA main electron beam for the FEL. The accelerator consists of 10 parallel beamlines. Each of them operates at a 10 kHz repetition rate. The beam can either drive 10 undulators or be combined into one 100 kHz beam to drive a single undulator. The drive beam to generate the high gradient (114 MV/m unloaded gradient) in the collinear wakefield accelerator runs at 325 MeV, 1.6 nC/bunch, and dual frequencies (3 MHz and 100 kHz). The drive beam uses a longitudinally triangular shape to balance the transformer ratio and gradient. The drive-to-main beam efficiency is close to 40%, which is expected to lead to a wall plug efficiency of the entire machine in the range of 5~10%.

In the remainder of this note, we discuss: the dielectric wakefield accelerator in Section II; the overall layout of the straw man collinear DWA to drive the future FEL light source in Section III; a few important issues including the shaped bunch generation for a high transformer ratio in Section IV; the design of the DWA structures in Section V; the beamloading estimation in Section VI; and cooling of the structure in Section VII. Finally, the note ends with a list of future work.

II. Collinear Dielectric Wakefield Accelerators

As one of the most promising techniques in the category of advanced accelerator concepts for high energy physics research applications, the wakefield accelerator is undergoing extensive investigation. In a collinear dielectric wakefield accelerator (CDWA), the fields generated by a leading, high-charge *drive* bunch (either a single drive bunch or a train of drive bunches) is used to accelerate a trailing, low-charge *main* bunch which contains a relatively small amount of charge. The collinear configuration has the two beams traversing the accelerator along the same trajectory as opposed to the Two-Beam Acceleration (TBA) scheme where the beams pass through different structures (decelerator and accelerator). In TBA, the wakefield energy is extracted from the decelerator to provide rf power for a conventional accelerator. On the other hand, the collinear wakefield accelerator directly transfers the energy from the drive bunch radiation to the trailing main bunch.

The Dielectric Wakefield Accelerator (DWA) is one of the major wakefield acceleration schemes studied over the past two decades. The most commonly used DWA structure is very simple - a cylindrical, dielectric tube with an axial vacuum channel is inserted into a metal guide (Fig. 1). The dielectric constant and the inner ("*a*") and outer ("*b*") radii of the dielectric tube are chosen to adjust the fundamental monopole mode (TM_{01}) frequency excited by the passing beam [11]. The phase velocity of the mode will equal the beam velocity ~c. When a high charge electron drive bunch travels through the vacuum channel of the DWA structure, it will lose energy due to electromagnetic Cherenkov radiation (wakefields) while propagating down the vacuum channel. Then a trailing, low-charge bunch at the correct delay can gain energy. The other DWA variations include rectangular DWA [12], multilayer DWA [13], and multi-channel DWA [14], etc.

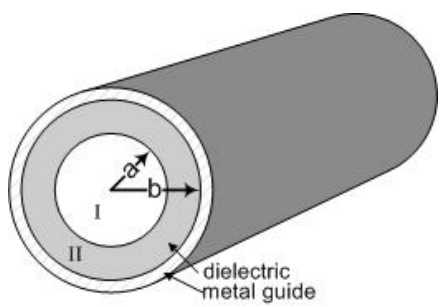


FIGURE 1. A Partially Dielectric-Loaded Circular Waveguide for particle acceleration. The dielectric tube has inner radius *a* and outer radius *b*. Region I is vacuum; region II is dielectric; and the outermost layer is metal.

A. Modes Analysis

In the application of particle acceleration, the monopole mode TM_{0n} mode (particularly TM_{01} for a single mode operation) is of main interest. For the TM_{0n} mode, the general solution for the electric and magnetic field components in this partial

dielectric-loaded circular waveguide can be derived from the wave equations with appropriate boundary conditions, which are shown in the following

$$E_z = \begin{cases} B_1 J_n(k_1 r) e^{j(\omega t - \beta z)} & 0 \le r < a \\ B_2 F_{nn} e^{j(\omega t - \beta z)} & a \le r \le b \end{cases}$$

$$E_r = \begin{cases} -\dfrac{j\beta}{k_1} B_1 J_n{}'(k_1 r) e^{j(\omega t - \beta z)} & 0 \le r < a \\ -\dfrac{j\beta}{k_2} B_2 F_{nn}{}'(k_2 r) e^{j(\omega t - \beta z)} & a < r \le b \end{cases} \quad . \tag{1}$$

$$H_\phi = \begin{cases} -\dfrac{j\omega\varepsilon_0}{k_1} B_1 J_n{}'(k_1 r) e^{j(\omega t - \beta z)} & 0 \le r < a \\ -\dfrac{j\omega\varepsilon_r\varepsilon_0}{k_2} B_2 F_{nn}{}'(k_2 r) e^{j(\omega t - \beta z)} & a < r \le b \end{cases}$$

B1, B2, A1 and A2 are the field amplitudes in the region I (vacuum) and II (dielectric) respectively, v_p is the phase velocity of the wave traveling inside the tube, and c is the speed of light in free space. k_1, k_2 are the cutoff wave numbers in the each region, and β is the wave propagation constant. They are defined as

$$k_1 = \omega \sqrt{\frac{1}{c^2} - \frac{1}{v_p{}^2}}$$

$$k_2 = \omega \sqrt{\frac{\varepsilon_r}{c^2} - \frac{1}{v_p{}^2}}$$

$$\beta^2 = k_0{}^2 - k_1{}^2 = \varepsilon_r k_0{}^2 - k_2{}^2$$

$$k_0{}^2 = \omega^2 \mu_0 \varepsilon_0$$

$$F_{nn}(k_2 r) = J_n(k_2 r) - \frac{J_n(k_2 b)}{Y_n(k_2 b)} Y_n(k_2 r)$$

$$F_{nn}{}'(k_2 r) = J_n{}'(k_2 r) - \frac{J_n(k_2 b)}{Y_n(k_2 b)} Y_n{}'(k_2 r) \, . \tag{2}$$

The electric fields in the vacuum region described by Equation (1) have very interesting characteristics. When the phase velocity of the wave is c, the transverse wave constant in vacuum region k_1=0 from Equation (2), which implies that E_z is constant in this region from Equation (1). This implies that there is no electric field enhancement on ceramic surface relative to the gradient on axis. This is critical for high gradient operation of the accelerating structure design. Figure 2 shows this characteristic of the longitudinal electric field for TM_{01} mode.

Using the boundary conditions at $r=a$ (E_z *and* H_ϕ continuous), the dispersion relation of guided wave traveling in this waveguide can be obtained as

$$\left[\frac{1}{k_1} \frac{J_n{}'(k_1 a)}{J_n(k_1 a)} - \frac{\varepsilon_r}{k_2} \frac{F_{nn}{}'(k_2 a)}{F_{nn}(k_2 a)} \right] = 0 \ . \tag{3}$$

From the Equation (2) we know the phase velocity of the guided wave in this waveguide can be slowed down below the speed of light due to the loaded ceramics. Meanwhile, the inner radius *a* is chosen based on the requirements of beam dynamics. Thus, governed by the mode dispersion relation (Equation (3)), the outer radius *b* should be adjusted accordingly to achieve the desired v_p.

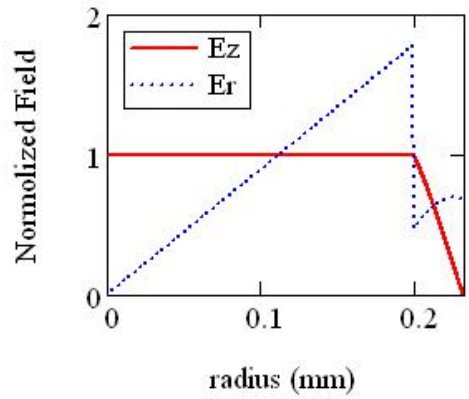


FIGURE 2. The amplitude of E_z and Er of the TM_{01} mode along the transverse radius of a dielectric-lined circular waveguide. The geometry of the waveguide is shown in Fig.1 with a=0.2 mm, ε_r=3.75, and b=0.232mm. The phase velocity of the TM_{01} wave is equal to c at 850GHz. Note that the relation of the maximum values of Er and Ez satisfies $E_{rmax}/E_{zmax}=(\pi a)/\lambda$

B. Analytical wakefield calculation

Normally, the calculation of wakefields is tedious and cumbersome. The complete solution to the inhomogeneous wave propagation equations is required. For wakefields in circular dielectric-loaded accelerating structures, a clear derivation and some results can be found in references [15] and [16]. Instead of finding Green's function directly, here we briefly summarize an approach to calculate the wakefield in DWA structures using intrinsic waveguide modes.

Rather than solving the bunch-excited fields through Maxwell's equations directly, the wakefields can be expressed as a sum over waveguide parameter *r*/Q, or equivalent of loss factor in most literatures (a function of the structure geometry only) for each mode [17, 18]. The mode selection for the wakefields calculation is based on the relation between the system response and signal excitation. The guided wave modes in the waveguide system selected to compute wakefields are under the criterion of the electromagnetic fields' patterns matching, i.e. mode matching between the excited wakefields mode (classified by the leading-order azimuthal harmonic) and the guided wave modes (waveguide intrinsic modes) that let them have similar field's space distributions. For example, TM_{0n} mode will match the longitudinal wakefield in the circular DWA structure. The transverse wakefields are also obtained from Panofsky-Wenzel theorem [19]. Equation (4) shows the relation of the loss factor of a point charged particle for each mode (k_{li}) and the *r* over Q for each guided wave mode at the accelerating waveguide:

$$k_{l_i} = \frac{\omega_i}{4}\left(\frac{r}{Q}\right)_i, \tag{4}$$

where, the subscript i indicates the ith mode, ω *is* the resonant frequency, and (r/Q) is the geometric dependent parameter while charged particles traveling on axis for each intrinsic mode (note here we only consider longitudinal wakefield), which is obtained through computing the stored energy U per unit length and accelerating electrical field E_z

$$\left[\frac{r}{Q}\right]_i = \left[\frac{|E_z|^2}{\omega U}\right]_i. \tag{5}$$

Then the longitudinal wake function (Green's function) can be simplified to

$$w_z = \sum_i 2k_{li}\cos(\omega_i t). \tag{6}$$

The summation of modes is an infinite series and it can be truncated by the finite bunch length of the drive beam, as in case of most wakefields calculations. Finally, the wakefield (wake potential per unit length) of an electron bunch in an accelerating structure, which has a current distribution of *I(t)*, can be calculated through a convolution integral

$$W_z(t) = -\int_{-\infty}^{t} I(\tau) w_z(t-\tau) d\tau. \tag{7}$$

In addition, it is necessary to take the group velocity effect into account in a wakefield calculation if the device is a high group velocity structure. Details can be found in [20].

C. Wakefield RF pulse in a Finite Length DWA structure

The wakefield calculation presented above is carried out with an assumption of the infinite long wakefield device. In this Subsection, we briefly describe the wakefield rf pulse generation in a finite length DWA structure which is very helpful later on when considering the power dissipation and cooling in the wakefield accelerator design.

Consider a charged particle beam (the drive bunch) traveling down the axis of a dielectric-loaded waveguide of length L. The drive bunch will excite a wakefield that radiates a traveling wave (TW) power, P_w, into the modes of the structure. Here, the electron bunches are always considered to be ultrarelativistic (the particle velocity equals c, the speed of light in free space).

Let a drive bunch traveling in the $z+$ direction enter the DWA structure of length L at the moment $t = 0$ and location $z = 0$ (Fig. 3a). At time t when the bunch is still inside the structure (Fig. 3b), the head of the RF pulse (moving at c) is located at $z = ct$, while the tail of the RF pulse (moving at v_g) is located at $z = v_g t$. When the bunch (and the head of the RF pulse) reaches the downstream end ($z = L$) of the structure (Fig. 3c)

the time is $t = L/c$ and the tail of the RF pulse has reached the position $z = v_g L/c$. At this moment, the generated RF pulse would begin to leak out of the structure (with assumption of a perfect match coupler or rf absorber) at $z = L$, but in this heuristic description we imagine the RF packet propagating out of the end of the structure at $z = L$. The tail of the RF pulse exits the structure (Fig. 3d) at time $t = L/v_g$. The duration of the RF pulse is the difference between the moment the bunch reaches the exit ($t = L/c$) and the moment the tail of the RF pulse reaches the exit ($t = v_g L/c$). Therefore, the RF pulse duration for a single-bunch excitation is given by

$$\tau_s = L(1-\beta_g)/v_g . \tag{8}$$

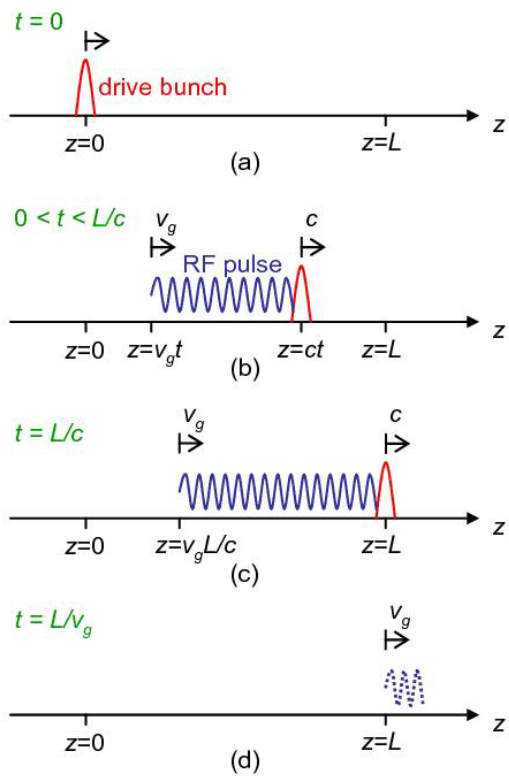


FIGURE 3. The RF pulse generated in the decelerator by a single drive bunch. Snapshots of a single bunch and the excited RF pulse inside the DWA structure region $0 \le z \le L$. (a) the bunch enters the structure ($t = 0$); (b) the bunch and the RF pulse are fully inside the structure and the head and the tail of the RF pulse travel at different speeds ($0 < t < L/c$); (c) the bunch reaches the exit of the structure ($t = L/c$); (d) the tail of the RF pulse exits the structure ($t = v_g L/c$).

Now we can calculate the power radiated by a drive bunch. Eqn.(6) and (7) gives the peak gradient of the wakefield for the ith mode generated by a single electron bunch can be expressed as

$$(G_b)_i = 2(k_l)_i q_b , \tag{9}$$

where q_b is the bunch charge.

Only considering the fundamental synchronous mode which is reasonable for properly designed structures with correction of the group velocity effect [20], equation (9) can be rewritten as

$$G_b = \frac{1}{2}\omega q_b \left(\frac{R}{Q}\right)\frac{F(\sigma)}{1-\beta_g}, \tag{10}$$

where $\beta_g = v_g/c$ is the normalized group velocity (c is the speed of light); and $F(\sigma)$ is the form factor of the bunch, which can be simply expressed as $F(\sigma) = \exp\left(-\frac{(k_z\sigma_z)^2}{2}\right)$ for the relativistic Gaussian bunch; k_z is the propagation constant of the excited mode and σ_z is the *r.m.s* bunch length.

Equation (10) is used to evaluate the maximum amplitude of the longitudinal wakefield signal (normalized to a 1 meter long structure) produced by a single electron bunch, and its output power can be calculated using the expression

$$P = \frac{\omega}{4}\left(\frac{R}{Q}\right)v_g \frac{q_b^{\,2}}{(1-\beta_g)^2} F^2(\sigma) e^{-2\alpha_0 v_g t} \qquad 0 \le t \le \tau_s\ , \tag{11}$$

where v_g is group velocity, and α_0 is the attenuation factor per unit length.

D. Wakefield Transformer Ratio vs. Gradient

In a wakefield accelerator, the fields generated by a leading, high-charge *drive* bunch (either a single drive bunch or a train of drive bunches) are used to accelerate a trailing, low-charge *witness* bunch. An important parameter that influences the performance of a wakefield accelerator is the transformer ratio $TR \equiv$(maximum energy gain of the witness bunch) / (maximum energy loss of the drive bunch). Therefore, to accelerate the witness beam to high energy it is desirable to make TR as large as possible. For instance, consider a plasma wakefield accelerator in which a 10 GeV drive bunch decelerating at a rate of 10 GeV/m (due to its self-wake) will be spent in 1 m. Over the same distance, the witness bunch will be accelerated by R·10 GeV. (The beam loading by the witness beam is neglected since we assume its charge to be low.) While TR is less than 2 under very general conditions: linear media; a relativistic, longitudinally symmetric drive bunch; and identical paths through the system of both drive and witness beams [21, 22], there are several techniques that can achieve TR > 2 by breaking the symmetry of the drive beam.

In addition to the transformer ratio, another central figure of merit in a collinear wakefield acceleration scheme is the accelerating gradient. The presence of a high accelerating gradient is always attractive because it will reduce the length of the main linac which is the primary determining factor for the facility size. However, these two parameters appear to have a complicated relation. Numerical simulations of a single bunch wakefield excitation reveal the dependence of the peak acceleration gradient (W^+) and TR on the bunch length σz (Fig. 4)[23]. As one would expect, the peak accelerating gradient increases as the bunch length decreases, but the behavior of TR as a function of

$\sigma z/\lambda$ (where λ represents the wavelength of the fundamental accelerating mode) is more complicated than the monotonically decreasing behavior of W^+. TR is peaked near $\sigma z=\lambda/4$, while dropping off towards 1 for bunch lengths on either side of $\lambda/4$.

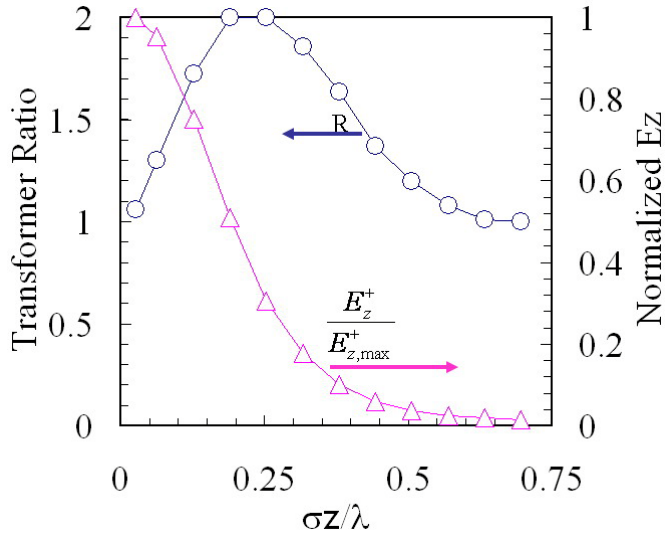


FIGURE 4. The transformer ratio as a function of bunch length. *TR* and peak accelerating field, E_z^+ , as a function of the normalized bunch length ($\sigma z/\lambda$). The plot is from Ref. [23]. (Single mode operation was assumed in the calculation.)

E. Single mode vs. Multimode

From Eqn.(7) the gradient in a DWA structure is determined by the peak current of electron bunch and the loss factor (equivalent to the impedance of the structure). The impedance of the structure strongly depends on the inner radius of DWA structure as well as the dielectric constant. A smaller beam channel yields higher impedance, but the effect of the dielectric constant is more complicated. In general, a lower dielectric constant leads to higher impedance. For the electron bunch with a given distribution, i.e. its frequency band is fixed; there exist two operational regimes to design a wakefield accelerator: single mode and multimode operations. In single mode operation, the structure is designed in such a way that only the lowest mode appears inside the bandwidth of the electron bunch. Figure 5 shows an example where the longitudinal wakefield and its spectrum are calculated for a beam of 1 nC and 50 micron bunch length passing through a quartz structure with inner radius of 0.2mm, outer radius of 0.232 mm. In multiple mode operation, just like an overmoded waveguide, multiple modes will be excited if we just increase the outer radius of the dielectric cylinder used to calculate the wakefield in Fig. 5; multimode results are shown in Fig. 6. Comparing Fig. 5 and Fig. 6, it is found that the wakefield gradient is the same for the two different structures. In fact, it is generally true that the wakefield gradient is the same for both single and multimode structure as long as their inner radius and dielectric constants are same. We note that a multimode structure may have an advantage regarding rf breakdown since it is exposed to high gradient for a shorter time than that in a single mode structure. On other hand, a single mode structure has the advantage regarding higher transformer ratio for the same

constrains. Lastly, both structures have similar rf losses. This is surprising at first since the high fields are present for a shorter time in the multimode structure, but this is compensated by its more higher order modes. Detailed discussions are beyond the scope of this note, but interested readers may refer [24] for more details.

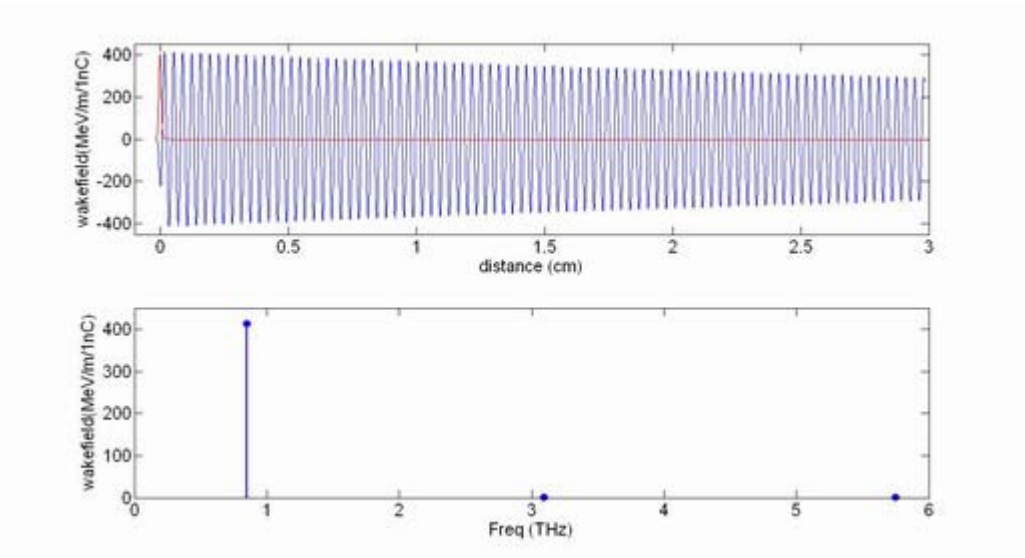


FIGURE 5. Single mode wakefield excitation due to a 1nC, 50micron of *r.m.s* bunch length electron bunch passing through the structure described in the caption of Fig.2.

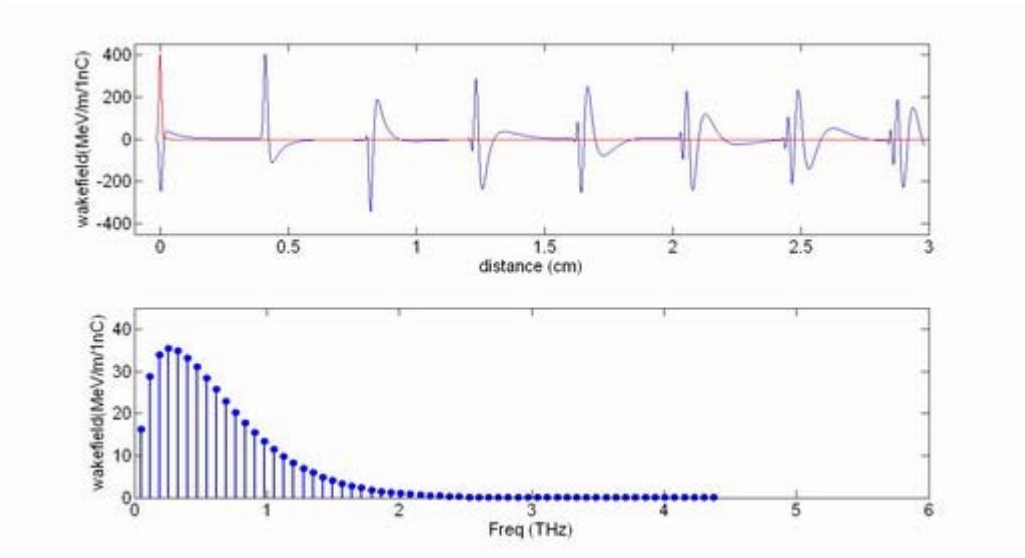


FIGURE 6. Multimode wakefield excitation due to a 1nC, 50micron of *r.m.s* bunch length electron bunch passing through the similar structure used in Fig.5 but a larger outer radius (1.42mm).

III. The Overall Layout of DWA for FEL

The overall layout is shown in Figure 7. Although at this very early stage all parameters are subject to change, some of them are slightly more rigid than others. For example, the repetition rate of the main beam is preferred to be in the >100 kHz range, charge in the range of 50~100 pC, and beam energy in 5~10 GeV range, etc. Also, some restrictions are on the drive beam side, like beam power and bunch length. The scheme we propose consists of 10 parallel main beamlines sharing one drive beam. Each beamline has two 50 m long 4 GeV accelerator modules. The 325 MeV drive beam is generated by L-band CW superconducting linacs. The drive beam has a dual frequency structure: the bunch train repetition rate is 100 kHz with microbunch frequency of 3 MHz (333.3 ns spacing) between two microbunches. The two microbunches drive the two accelerator modules of each beamline sequentially. Meanwhile, the drive beam distributes every 10th bunch to drive the ten beamlines in order so that the repetition rate for each beamline is one tenth of the drive beam frequency, i.e. 10 kHz. The main reason behind the choice of beam repetition rate is to maintain the average rf power dissipation of the accelerating structure at a level that conventional water cooling technology can handle (advanced water cooling techniques may be required though, see Section VII). The drive bunch has a longitudinally shaped profile for the purpose of increasing the transformer ratio (details will be discussed in Section IV). This double triangular beam can be generated using an emittance exchanger (EEX) at ~100 MeV level. The technique is under investigation and it is planned to conduct an experiment at AWA [25]. Simulation results are very encouraging at this early stage.

Each accelerator beamline consists of two modules. Each provides 4 GeV gain for the main beam. The total main linac length is 100 m, so that the effective gradient of the linac is 80 MeV/m, much higher than the conventional technologies in the CW regime. The total length of one module is 50 m, which matches the time spacing between two drive microbunches (333 ns). Before being injected into the collinear wakefield accelerator beamlines, the drive beam pulses are distributed along a common transport line in parallel with but traveling in an opposite direction to the main beam so that the distance between micro drive bunches has to be twice the length of one accelerating structure to ensure the arrival time at the entrance of the wakefield accelerator coincides with the main beam. The effective acceleration length in one module is 40 m. This is the distance traveled where the drive beam has lost 85% of its energy. An 80% fill factor is a reasonable number considering the beam optics and diagnostics in the beamline. In terms of beam breakup control, particularly for the dielectric wakefield accelerators (DWA), quadrupole channels are the most effective approach [26]. The approach used in the CLIC design to deal with the constantly increasing energy spread and emittance, i.e. quads plus BPM between structures, can also be adapted to a more advanced version of this scheme.

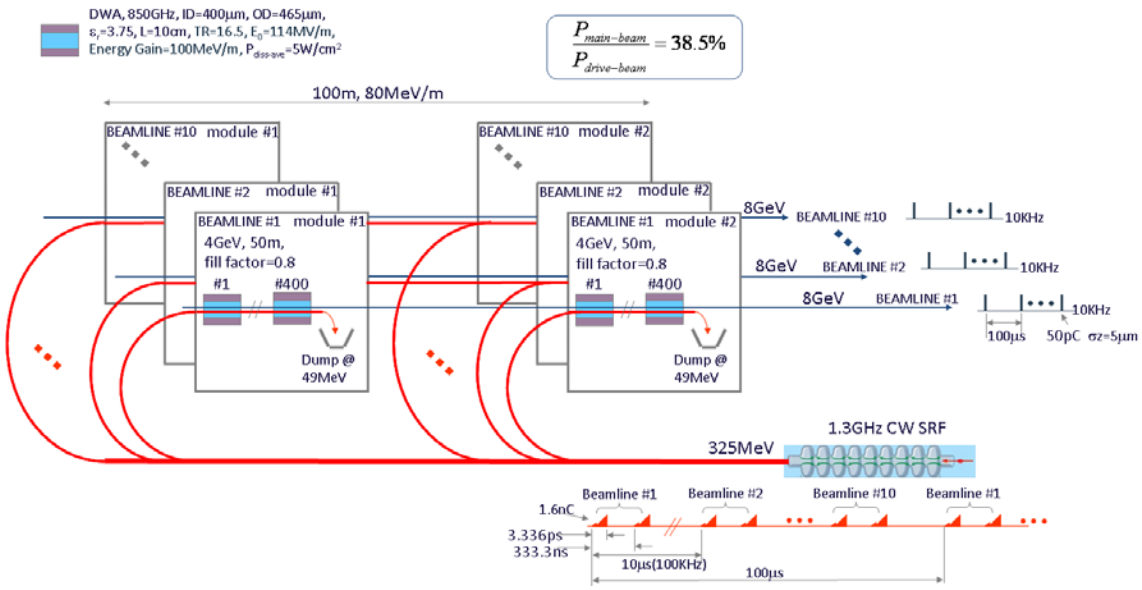


FIGURE 7. The overall layout of the electron beams in a dielectric wakefield accelerator based FEL light source facilty.

The layout shows the 400 10-cm long DWA structures in one module. The length of the structure is chosen for easy calculation. In reality, the structure can be any length insofar as the beam control requirement is satisfied. It should be pointed out that, in this single drive bunch wakefield acceleration regime, the rf power level has nothing to do with the structure length. It can be seen as an attenuated rf packet moving after the drive bunch at the group velocity (Dispersion may exist due to multimode generation). The peak value of the rf signal is set by superposition of the wakefields from the drive bunch and witness bunch (this beam loading effect is discussed in Section VI). The DWA structure has 400 micron beam hole, which is considered to be challenge but doable for a 325 MeV 1.6 nC bunch. The dielectric material to establish the Cherenkov radiation is quartz, which will reduce the structure cost considerably. Thermal heating due to the average dissipated power in quartz tube is quite small. Most of the heat comes from the surface current on the copper wall (the quartz dielectric loss is very small relative to the copper loss), which is estimated to be 5 W/cm^2.

The transformer ratio, defined as the ratio of the maximum accelerating field behind the drive over the maximum decelerating field inside the bunch, is 16.5 in our straw man design. The peak gradient behind the bunch is 114 MV/m which means the drive bunch energy loss is 6.9 MeV/m. All the values are calculated with four longitudinal modes included, but a considerable contribution is from the fundamental mode because the structure was intentionally designed to minimize the loss factor of all high order modes (HOM). This is very important for the drive bunch to achieve a high transformer ratio, as well as to mitigate the self wake inside the main beam (the main beam has much shorter bunch length which is prone to excite HOMs). Detailed structure design will be discussed in Section V.

Two major factors were considered in the choice of the main bunch parameters: 1) to minimize beam loading the charge per bunch was made as high as possible without

causing serious drop in the net gradient; 2) to maximize beam brightness we kept the charge as low as possible to yield small energy spread and a short bunch/high peak current. A short bunch makes it easier to minimize the energy spread. The self wake from the main beam depends on the wakefield structure, is proportional to the charge, and inversely proportional to the bunch length. In the collinear wakefield acceleration scheme, both the drive and main beam share the same beam channel. In general, while the structure is optimized to maintain a single mode excitation for drive beam, it will be a multimode excitation for the main beam because of its shorter bunch length. But the self wake inside the main beam will converge very fast because loss factor of HOMs become negligible after the first few modes. 50 pC/ bunch and 5 μm bunch length (*r.m.s.*) are chosen in the design is to allow a reasonable beam loading and ~1.2 kA peak current. As shown in Figure 7, the energy gain per unit length of the main beam still remains 100 MeV/m (centroid energy). The associated energy spread is ~1.5% (see Section VI).

Another critical parameter of an accelerator is the efficiency. The overall efficiency of the proposed scheme can be roughly estimated by cascading the efficiencies of each subsystem, for example, klystron AC-to-RF efficiency, RF-to-drive beam efficiency, and drive-to-main beam efficiency, etc. The drive-to-main beam efficiency, 38.5%, can be obtained using the parameters in the layout. If we assume 50% efficiency for each of other subsystems, the overall wall plug efficiency is around 5~10% with consideration of power consumption from the cooling, magnets, and infrastructure, etc.

-----*An Alternative Scenario*

Two central figures of merit in a collinear wakefield acceleration scheme are the transformer ratio and the accelerating gradient. The former affects the energy transfer efficiency from the drive beam to the accelerated witness beam and is eventually linked to the overall efficiency (under some assumptions, e.g. fixed beam loading, the same wakefield device, etc.). Meanwhile the presence of a high accelerating gradient is attractive because it will reduce the main linac length which is the primary determining factor for the facility size. The general relation of the transformer ratio and accelerating gradient has been discussed in Section II.

A trade off must be made between the transformer ratio and the accelerating gradient. To demonstrate how this tradeoff can be made we calculate the gradient and transformer ratio for the structure described in Fig.7 driven by both a Gaussian and a triangular drive bunch of 1 nC and 0.5 ps FWHM. The Gaussian bunch generates ~325 MV/m with a transformer ratio of 2 which means that the drive beam will be decelerated by 16.25 MeV in a 0.1 m long structure. Moreover, it means that the energy of the drive beam will be spent after traveling 2 m and the energy gain of the trailing main beam will be 650 MeV. Using the same the modulator length (50 m) as shown in Fig.7 (this length was determined by drive bunch train spacing of 3 MHz), the fill factor is 2/50 so that the effective gradient drops to ~13 MV/m. To increase the effective gradient the transformer ratio can be made larger than 2. However, previously studies [25] have shown that when the pulse is shaped to increase the transformer ratio, the wakefield gradient is lowered. For the case shown in Fig. 7, the transformer ratio is 16.5 and the wakefield gradient decreased to 114 MV/m. If we use a 100 pC/bunch main beam

instead of 50 pC in Fig.7, the energy gain of the main beam will be reduced to ~ 86 MeV/m approaching gradients achievable by conventional systems.

Transformer ratio enhancement using a temporally shaped bunch has yet been demonstrated. On the other hand, transformer ratio enhancement using the ramped bunch train technique has been demonstrated in two experiments performed jointly by Euclid and ANL/AWA. The transformer ratio was enhanced to R=2.3 using a ramped bunch train of 2 bunches in a 13.65 GHz dielectric-loaded accelerating structure and was first observed in 2006 [27]. Later, R=3.4 was achieved in a second experiment with the improvement of matching the bunch length (lengthened from σ_z=2 mm to 2.7 mm using a laser stacking technique) with the wavelength of the fundamental mode of the same wakefield device [28]. In the case of using a photocathode gun to provide the beam source, the RBT scheme is easier to implement compared to the bunch shaping scheme because the laser beam splitting and intensity control is done outside of vacuum. In Fig.8, we provide a similar configuration to the previous figure (Fig. 7) but using an RBT to enhance the transformer ratio instead. Four Gaussian bunches with total charge of 1.6 nC, which generate a 131 MV/m gradient in the same structure used in Fig.7, are the replacement of one double triangular bunch used in Fig.7. Since the transformer ratio is only half of the previous one, each module can only provide a 2 GeV gain for the main beam. Therefore four accelerator modules are needed for each beamline to boost the main beam energy to 8 GeV. The microbunch repetition rate of the drive beam is increased to 6 MHz to improve the fill factor of each module so that the effective acceleration remains the same as the first configuration, 80 MeV/m. Under the same main beam current, the loaded gradient in the RBT scenario is 20% higher so that the total structure length (or structure numbers) can be made smaller and, therefore, the structure cost will be lower. But the price paid is a decrease of the efficiency. This is a good example to illustrate the complexity of balancing the transformer ratio and gradient.

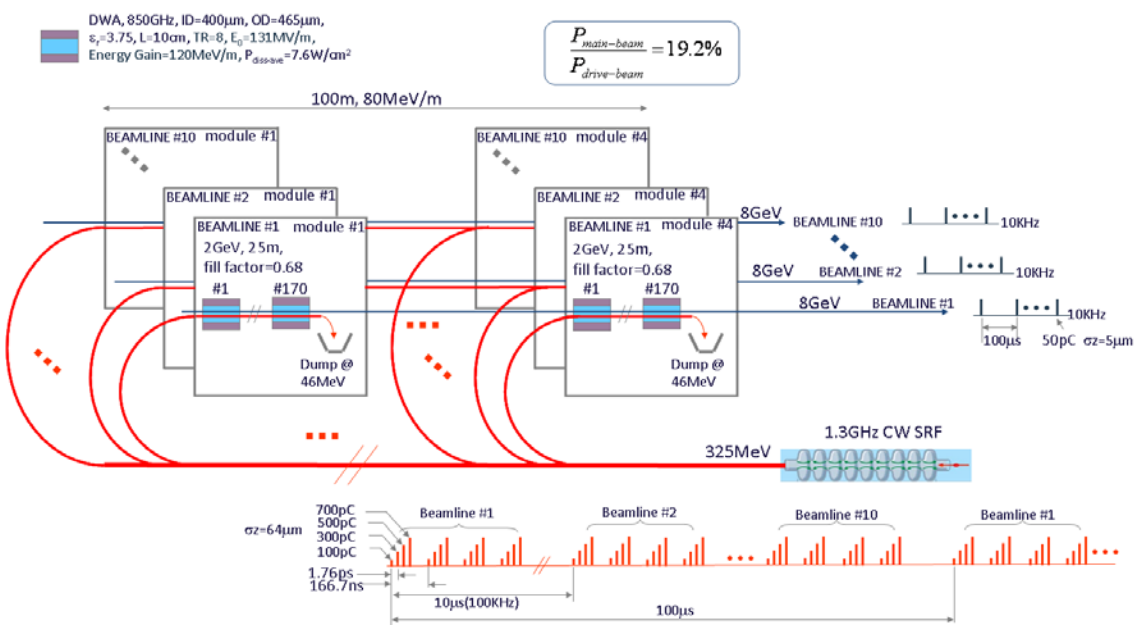


FIGURE 8. An alternative layout using a ramped bunch train technique to enhance transformer ratio. The RBT technique has been demonstrated experimentally at the AWA.

IV. Enhance Transformer Ratio Using A Shaped Bunch

For a finite length, longitudinal symmetric bunch, the transformer ratio TR can never exceed 2 [21].The general solution to increase the transformer ratio TR above the ordinary limit of 2 in a collinear wakefield accelerator is to use a bunch or bunch train which has an asymmetric temporal profile. For example, a bunch with ramped charge (*triangular beam and some variations*), a ramped bunch train (*a train of electron bunches with specific ramped-up charge ratio but equal time separation, i.e. amplitude modulated bunch train*), and a "compressed" bunch train (*a train of electron bunches with specific time separation among bunches but equal charge, i.e. a spacing modulated bunch train*), etc. The first two approaches are more attractive because they can enhance the transformer ratio, *T*, more rapidly. For example, in the ideal case, transformer ratio of a triangular bunch is $T \approx 2\pi N$, where N is ratio of the total bunch length over wavelength of the generated wakefield (single mode excitation). In the same duration, 2N micro-bunches can be arranged with half wavelength spacing between microbunches and charge ramped up with ratio of odd numbers (1:3:5:7…). Then this ramped microbunch train can raise the transformer ratio to $T=4N$ with an assumption that the single bunch transformer ratio is 2.

A. Bunch Shaping to Enhance the Transformer Ratio

Another approach to enhancing the transformer ratio is to use a triangular bunch or one of its variations (e.g. a "doorstep" bunch, a constant charge distribution for the first quarter of wakefield wavelength then followed by a triangular profile), which is particularly attractive for applications where only a single drive bunch is needed. Various methods to generate such a beam profile have been explored including compressing a time-energy chirped bunch with an anisochronous dogleg beam line [29] and by temporal shaping of the drive-laser intensity [30]. However, each of these methods has limitations. The techniques based on shaping of the photocathode drive laser distribution are prone to the space-charge effects which are prominent at low energies and tend to smear the impressed shaping. Similarly the techniques based on nonlinear transformations generally have a limited tunability.

Lately, a new variation of triangular bunch---double triangular bunch is proposed to enhance the transformer ratio [25] (it is used in the scheme shown in Fig. 7). The current profile of the double triangular bunch can be described as

$$I(t) = \begin{cases} I_0 \omega t & 0 < t < \dfrac{\pi}{2\omega} \\ I_0 \omega t - I_0 & \dfrac{\pi}{2\omega} < t < T \end{cases}, \tag{12}$$

where ω *is* the resonant frequency of the wake function, *T* is the bunch duration, I_0 is the total charge related constant current. Figure 9 illustrates current profile of the double triangular bunch and its corresponding longitudinal wakefield generated in a single mode wakefield accelerator. The maximum achievable transformer ratio in theory is

$$R = \sqrt{1 + (\omega T - 1)^2}\,, \tag{13}$$

where, ω is the resonant frequency of the wake function and T represents the bunch duration. For example, transformer ratio TR is 17 in the structure (850 GHz) shown in Fig.7 if the bunch duration is 1 mm.

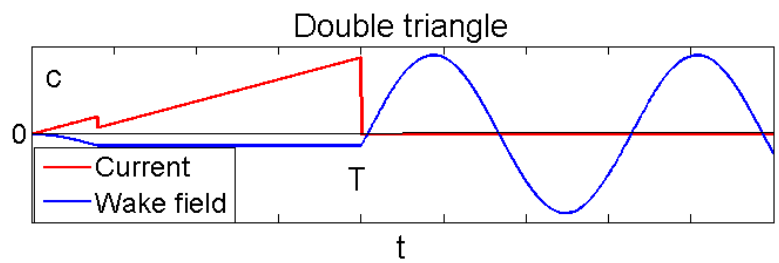


FIGURE 9. The current distribution of the double triangular bunch and its wakefield in a single mode wakefield device.

Influence of a slowly dropping tail In our analysis, the current density distribution of the triangular bunch, or its variations, is assumed to have a steep cutoff edge at the bunch tail. In reality, space charge effects, coherent synchrotron radiation, and energy divergence of the bunch will diffuse the particles in phase space and cause some difficulties in generating such a sharp edged bunch during experiments. To address this issue, consider the current distribution shown in Fig.10. For example, if $T_2=2T_1$, the distribution will return to a nearly symmetric one and the transformer ratio will decrease to around 2. To minimize this tail effect, it requires that all the particles at the tail (particles appear in $T_1<t<T_2$) slip away from the accelerating phase of the wakefield, i.e., they all lose energy and contribute to the wakefield behind. Meanwhile, the decelerating field of the tail should not be greater than that from the front of bunch. These conditions can be expressed as follows,

$$\begin{cases} V^-(0 < t < T_2) < 0 \\ \left|V^-(0 < t < T_1)\right| \geq \left|V^-(T_1 < t < T_2)\right| \end{cases}, \tag{14}$$

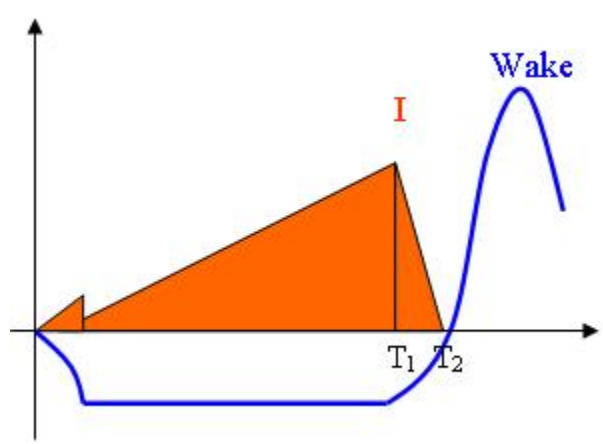


FIGURE 10. The double triangular bunch with a dropping tail and its associated wakefield.

With assumption of a linear tail from T1 to T2, the condition to maintain the unaffected transformer ratio is

$$\cos(\omega\Delta T) \geq \frac{\omega T_1 - 1}{\omega T_1 - 1 + \omega\Delta T}, \tag{15}$$

where $\Delta T = T_2 - T_1$. For example, if $\omega T_1 = 4\pi$, then $\omega\Delta T < 0.054\pi$ should be satisfied to have the maximum transformer ratio of 3.7π.

B. Shaped Bunch Generation using Emittance Exchanger

The reason to choose the double triangular bunch is that it is relatively easy to implement using a new approach: bunch shaping via an emittance exchange (EEX) beamline [25, 31].The EEX beamline, in a general sense, can be treated as a transport matrix between the transverse and longitudinal phase space, in such a way that one can manipulate the beam for different needs. For example, if a specific transverse mask is applied at the entrance of the EEX beam line then a longitudinal distribution can be shaped accordingly at the beamline exit. Figure 11 shows the sketch of a typical EEX beamline, which is composed of two dispersive sections (so-called doglegs) with an RF deflecting cavity in between.

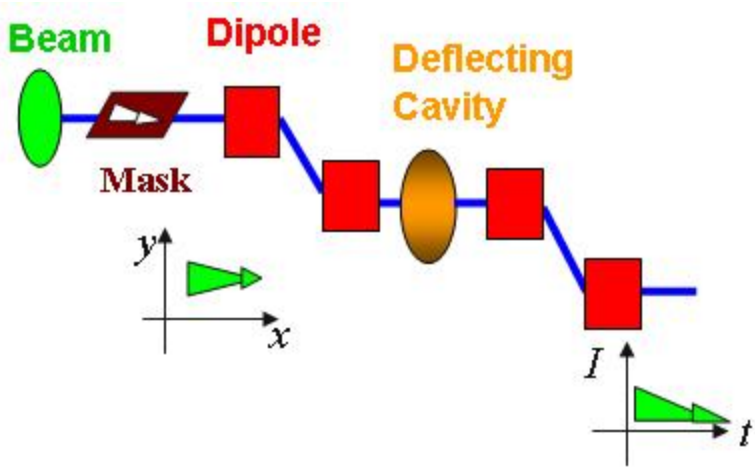


FIGURE 11. Simplified sketch of athe EEX beamline.

The physics behind of EEX beamline has been well presented in [32, 33]. Here we just introduce it in a very simple manner. The EEX beam line is composed of two dispersive sections (so-called doglegs) and a RF deflecting cavity in between. The transport matrix of the dogleg M_D can be expressed as (only the horizontal (x) and longitudinal (z) planes are considered) [33]:

$$M_D(\eta, \xi, L) = \begin{bmatrix} 1 & L & 0 & \eta \\ 0 & 1 & 0 & 0 \\ 0 & \eta & 1 & \xi \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{16}$$

where L, η, ξ are the length, dispersion and momentum compaction of the dogleg respectively. The RF deflecting cavity, which operates at the TM_{110}-like mode, delivers a longitudinal electric kick to the particle with strength in proportion to the distance from the axis. The transport matrix M_c of deflecting cavity for thick lens approach can be written as follow [34]:

$$M_C(k) = \begin{bmatrix} 1 & L_c & kL_c/2 & 0 \\ 0 & 1 & k & 0 \\ 0 & 0 & 1 & 0 \\ k & kL_c/2 & k^2L_c/4 & 1 \end{bmatrix}, \tag{17}$$

where L_c is the length of the cavity, $k=V_0/(aE_0)$ is the cavity strength parameter, V_0 is the cavity voltage, $a=\lambda/(2\pi)$, λ is the wavelength of the TM_{110}–like mode, and E_0 is the beam energy. For a complete exchange, the dispersion η and the cavity strength k should satisfy [33, 35]:

$$1 + k\eta = 0. \tag{18}$$

With a combination of Eq. (16)-(18), the coordinates of a particle at the exit of the EEX section can be obtained:

$$\begin{bmatrix} x_2 \\ x_2' \\ z_2 \\ \delta_2 \end{bmatrix} = \begin{bmatrix} 0 & L_c/4 & (\frac{L_c}{4}+L)k & -\frac{1}{k}+(\frac{L_c}{4}+L)k\xi \\ 0 & 0 & k & k\xi \\ k\xi & -\frac{1}{k}+(\frac{L_c}{4}+L)k\xi & \frac{L_c k^2\xi}{4} & \frac{L_c k^2\xi^2}{4} \\ k & (\frac{L_c}{4}+L)k & \frac{L_c k^2}{4} & \frac{L_c k^2\xi}{4} \end{bmatrix} \begin{bmatrix} x_1 \\ x_1' \\ z_1 \\ \delta_1 \end{bmatrix}, \tag{19}$$

where (x_i, x_i', z_i, δ_i) denotes coordinates of the particle at horizontal and longitudinal phase planes, and the subscripts 1, 2 represent that before and after EEX section respectively. For a thin cavity (L_c=0), the transform matrix in Eqn. (19) will be reduced to block anti-diagonal so that the transverse and longitudinal coordinates are directly mapped to each other. Typically, L is about a meter, L_c is about 0.1 m, ξ is about several millimeters, and k is usually less than 10.

An EEX beamline will be installed in Argonne Wakefield Accelerator (AWA) facility. It is a great opportunity to use this platform to demonstrate high transformer ratio with a double triangular bunch in a near future.

The preliminary beam simulation has been initiated but the main optimization work will be performed in the near future. The start to end simulation of the double triangular bunch generation was carried out using PARMELA. The space charge effects were included in the simulation. The mask and the deflecting cavity were applied through a homemade Matlab® routine. The injected beam was created using the AWA photocathode gun [36]. As examples, here we briefly present preliminary results of two simulations. One has a shorter bunch length (~1 mm) in order to maintain a relatively high gradient while increasing the transformer ratio, but a ~400 GHz structure is required to achieve R~5; the other has a longer bunch length (20 mm) so that the high transformer

ratio can be achieved in a 'low' frequency DWA structure (~30 GHz). It should point out that, with using the high brightness beam (i.e. not bounded with the AWA beam), the transformer ratio can be tuned much higher while relatively high gradient maintained.

Case I---a short shaped bunch:

The injected beam parameters are listed in Table I, and the EEX beam line settings in Table II. As a short bunch is desired for high wake field gradient, the EEX is performed at 75 MeV because the space charge effect is small at this level and the un-normalized transverse emittance is reduced as well. Both of these conditions favor creating a short bunch after the EEX.

TABLE I. AWA injection beam parameters

Beam energy	75 MeV
Bunch charge	3 nC
Transverse emittance (normalized)	19 mm mrad
Energy divergence (correlated)	558 keV
Energy divergence (uncorrelated)	7 keV
Bunch length	1 mm

TABLE II. EEX beam line settings

Dipole length	0.327 m
Bending angle	22º
Drift space Between bands	1.0 m
Deflecting cavity strength k	2.9

A chicane was designed for use upstream of the EEX beamline (see Fig. 12) to force the condition on the incoming longitudinal phase space: $d\delta / dz \approx -1/\xi$ [33]. This condition reduces the emittance dilution through the EEX beamline. A tunable collimator is installed in the middle of the chicane to chop out the diverging particles which may be difficult to control in the EEX. The beam at the entrance to the mask (after the chicane and the energy chopper) is shown as blue section of the curve in Fig. 13a. The beam at the entrance to the exchanger (after the transverse mask) is shown in Fig. 13b. The beam charge after the mask is 0.92 nC. At the exit of the EEX beamline, a longitudinally shaped beam is obtained (see Fig. 14). The bunch length is about 1 mm. The transformer ratio for an ideal 400 GHz single mode structure driven by this beam is 5. Note that the final longitudinal phase space distribution is not quite a double triangle in this simulation. More simulation work will be done in the near future. Nevertheless, this shaped bunch creates a relatively flat self wake, which is the key factor for increasing the transformer ratio. The normalized transverse emittance of the output beam is 36 mm mrad. After the quad, the beam size is focused to +/- 0.4 mm, where the dielectric wake field tube is placed.

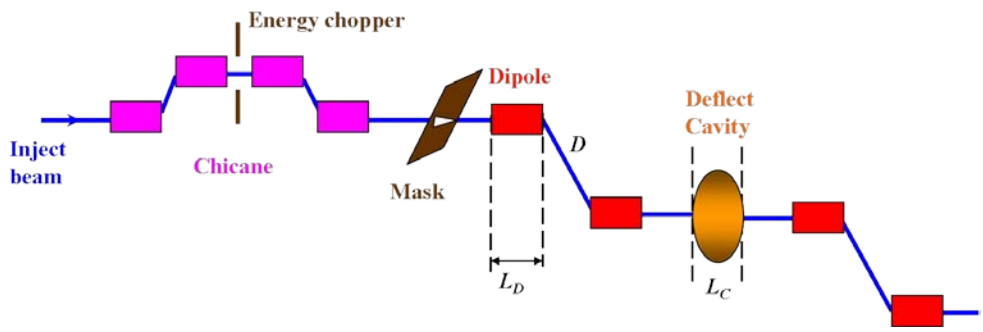


FIGURE 12. Sketch of the EEX beam line to generate a ramped bunch.

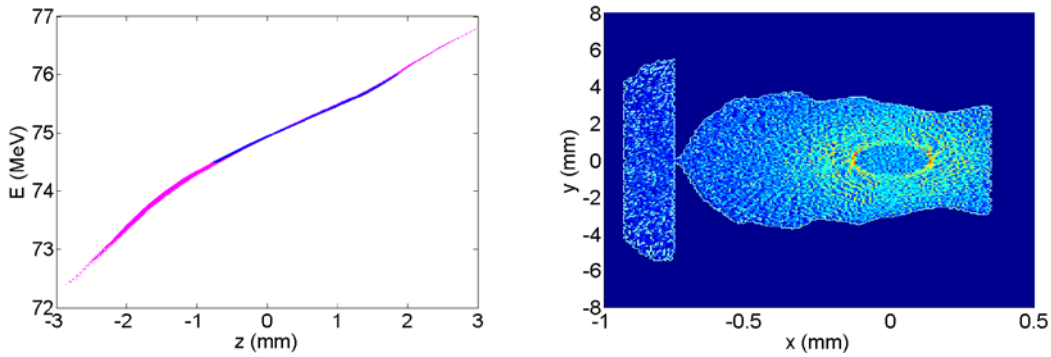


FIGURE 13. (a) Longitudinal phase space of the beam after a chicane and energy chopper (blue ones remain); (b) Transverse shape of the beam after the transverse mask.

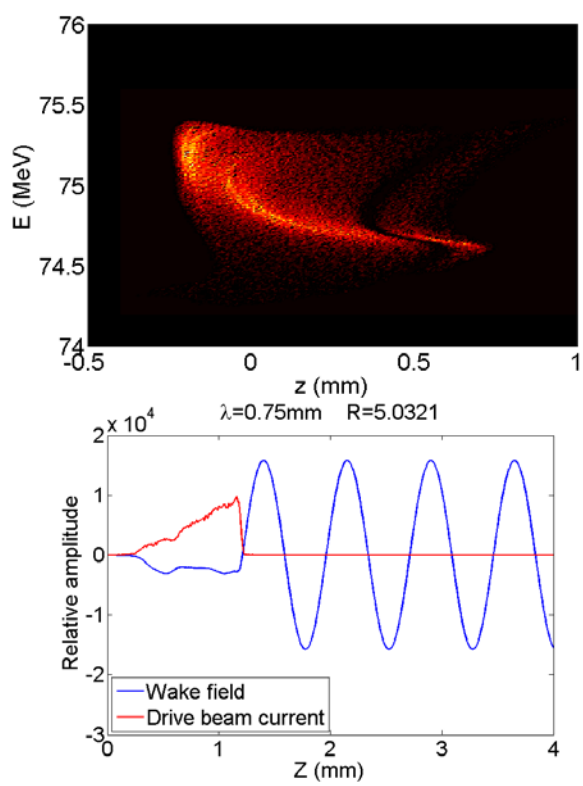


FIGURE 14. (a) The final beam longitudinal phase space; (b) Longitudinal beam distribution and its wake field in a 400 GHz single mode device.

Case II---a long shaped bunch at the beam energy 30 MeV:

In order to demonstrate a high transformer ratio in a microwave range wakefield device, a longer shaped bunch is needed. Table III shows the major beam parameters before the transverse mask and the EEX beamline settings to create said bunch. Following the same procedure, a nearly perfect double triangular bunch is obtained. Figure 10 shows its longitudinal shape and the wake generated in a ~32GHz wakefield device. *R~8* is obtained.

TABLE III. beam parameters before the EEX mask and the EEX beam line settings

Bunch charge	30 nC
Beam energy	30 MeV
Beam size (RMS x\y)	0.7 cm\0.3 cm
Bunch length (RMS)	1.7 mm
Energy divergence (RMS)	95 keV
EEX line dipole length	0.5 m
Bending angle	45º
Deflect cavity voltage	3.6 MV (peak)
Drift space (between 2 dipoles)	0.3 m

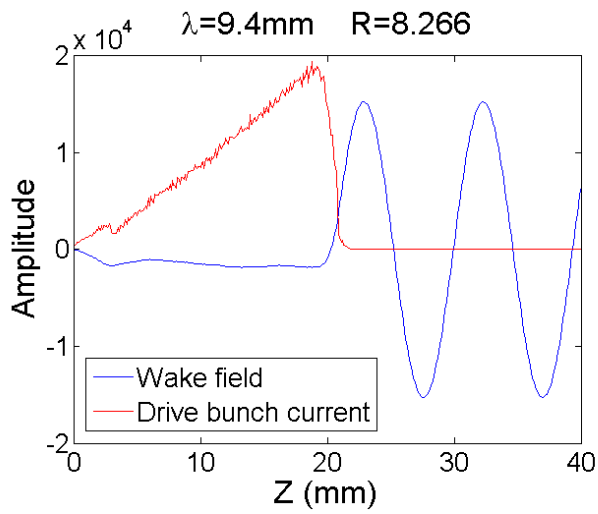


FIGURE 15. (simulation) Ramped bunch distribution created by the EEX beam line and its wakefield.

Note that it is impossible to generate an ideal double triangular bunch. However, in order to explain the fundamental concept of the DWA scheme mostly clearly we use the ideal shaped bunch for the wakefield calculation, structure design, and other discussions. More realistic design and parameters will be studied while moving this conceptual layout forward.

V. Considerations of a Single Mode DWA: Design and Wakefield Calculation

Due to the simplicity, low fabrication cost, and potential for high gradient we chose a dielectric wakefield accelerator (DWA) for the main linac. A DWA can be as simple as a dielectric tube surrounded by a metal jacket. The structure characteristics are determined by the 3 DWA parameters: inner radius of the dielectric tube a, outer radius b, and dielectric constant ε. In practice, a is chosen first to be consistent with the transverse size of the beam. The wakefield strength (we are only discussing the longitudinal wakefield here) is inversely proportional to the inner radius a. However, considering the beam loading effect, a cannot be too small because both drive and main beam share the same beam channel. The frequency spectrum of the intrinsic waveguide modes are discrete and go to infinity and can be adjusted by the proper choice of ε and b. On the other hand, the frequency spectrum of the beam is continuous and bandwidth limited. The bandwidth of the beam is determined by the temporal shape of the bunch, for example, the shorter a Gaussian bunch is, the broader its bandwidth. The waveguide spectrum excited by the beam is therefore only those covered inside the beam spectrum.

We have designed the DWA to operate in a single mode since theoretical considerations show that this choice maximizes the transformer ratio [37]. In order to implement the single mode operation, we can tune the 3 DWA parameters of the waveguide to move all high order modes (HOM) out of the drive beam bandwidth. For a DWA, this is done by varying the thickness of dielectric wall (thinner dielectric wall causes the waveguide modes to have a larger separation) and the dielectric constant (see Fig.16). In addition, these two knobs will also change the loss factor of each mode. Ideally, we can reduce the loss factor of the HOMs so that those modes will be excited very weakly. This point is particularly important as it offers a way of mitigating the self wake inside the main beam which has a wider bandwidth than the drive bunch since it is shorter than the drive bunch. The fast damped loss factor of HOMs will ensure the self wake of the main beam converged fast so that the loaded gradient will not drop seriously even for an extremely short main beam. Finally, the operating frequency of DWA structure (ω) is linked to a, b, and ε through the dispersion relation of the DWA structure, and it determines the transformer ratio of a shaped drive bunch of duration T (see Eqn. (13)).

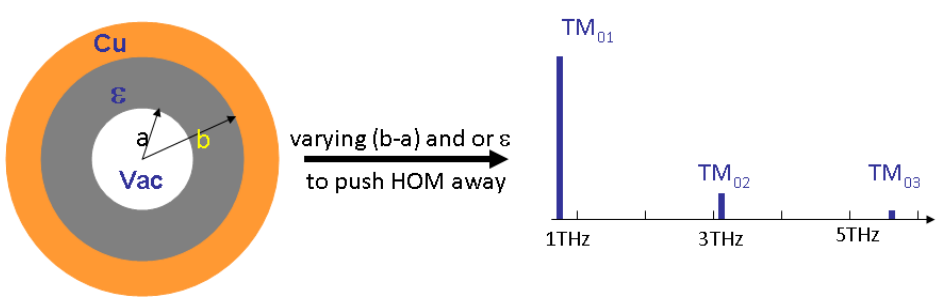


FIGURE 16. By tuning the parameters of the DWA, the HOMs are pushed out of the bandwidth of the drive beam so that only a single mode is excited. In addition, the loss factor of the HOM's are heavily damped to reduce the self wake of the much shorter main beam.

Bunch Properties

Figure 17 shows the double triangular bunch's longitudinal shape and its corresponding spectrum (solid red lines) that we used in our straw man study. For comparison, a Gaussian bunch (dotted blue line) is plotted as well. Unlike the Gaussian bunch, we see that the spectrum of the triangular bunch has a sharper drop at the beginning but then rolls off more slowly above the frequency ~1/*T*. This feature causes difficulties for single mode operation of a DWA driven by a shaped bunch.

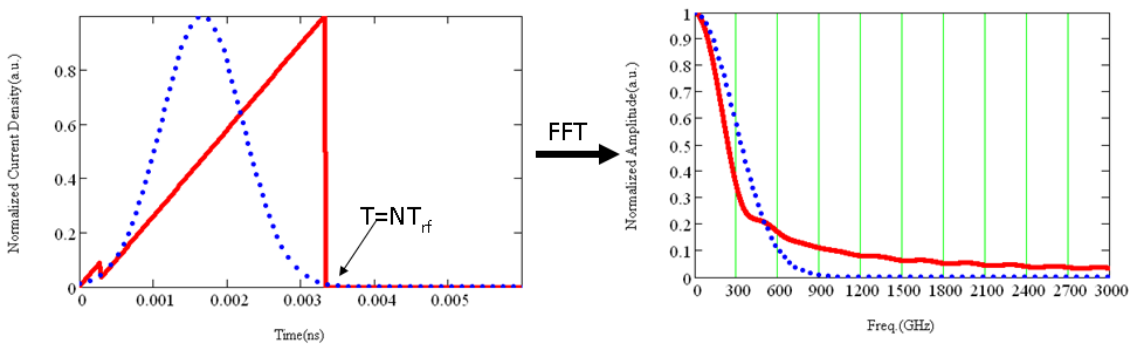


FIGURE 17. Comparison of a double triangular to a Gaussian bunch with $6\sigma_t=T$, where σ_t is the bunch length of the Gaussian bunch and T is total bunch length of the double triangular bunch.

Ideally, we would like to operate the DWA with high transformer ratio and high gradient. However, as stated above, choosing a shorter bunch will excite a stronger wakefield but decrease the transformer ratio and vice versa. Alternatively, reducing the radius of the DWA, *a*, will increase the gradient while not directly affecting the transforar ratio. Unfortunately, this approach is not without constraints either since the radius, *a*, cannot be smaller that the drive beam. Furthermore, the self wake of the main beam ("beam loading") will increase even more rapidly than the accelerating wakefield of the drive beam since it is much shorter. Taking the above into account our design philosophy was to achieve high TR (>>*2*) while maintaining a moderate gradient so that the energy gain of the witness beam is no less than 100 MeV/m.

Let us consider the dependence of the transformer ratio and gradient upon the bunch profile. We first assume that the double triangular bunch spans from time zero to *T*, and *T* satisifies $T=N*T_{rf}$, where T_{rf} is the period of the wakefield behind the bunch and *N* is a positive real number. Again, here we only consider the single mode scenario. The average current inside the drive bunch, I_{ave}, can be obtained through integration of its current density distribution described in Eqn. (12):

$$\left(\pi N+\frac{1}{4N}-1\right)I_0=\frac{Q_{total}}{T}=I_{ave}\,, \tag{20}$$

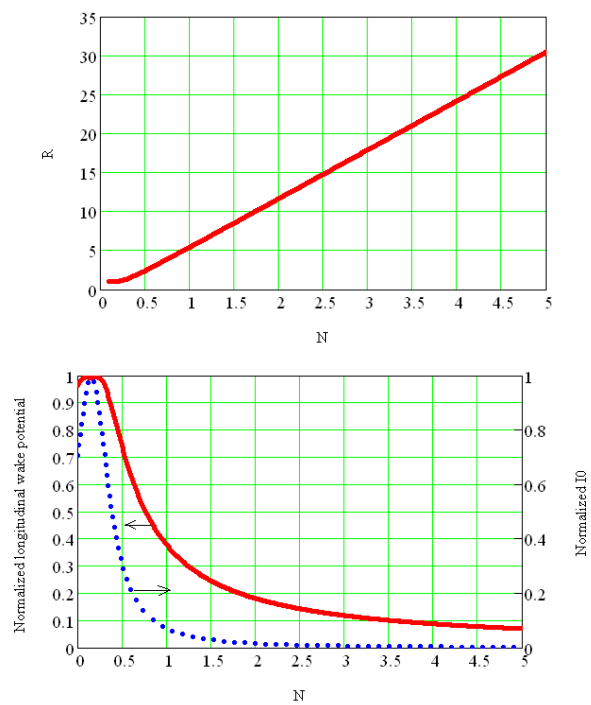
where Q_{total}, represents the total charge of a double triangular bunch. Following the approach presented in Section II (Eqn. (4), (6) and (7); with the current distribution in Eqn. (12)) we have peak wake potential of the drive beam

$$\left|V^{+}\right|=\frac{I_0}{2}\frac{R}{Q}\sqrt{(2\pi N-1)^2+1}\,, \tag{21}$$

and the transformer ratio

$$TR = \sqrt{1 + (2\pi N - 1)^2} \ . \qquad \text{(22, same as Eqn.(9))}$$

Once the bunch shape is given, Eqn. (22) clearly shows that TR is purely a function of N, and it is linearly proportional to N as long as N is not very small (see Fig.18(a)). This clearly implies that TR can be increased by making a long bunch (>>N). However, the wakefield gradient has a more complicated dependence on N (see Fig.18(b)). This is mainly because I_0 peaks at a small N but sharply deceases while N increases, which limits the positive influence from operating at large N. Note that the drive bunch charge was fixed in this parameter sweep, i.e. Q_{total} is a constant, and we assume R/Q of the structure does not change for different frequencies which is generally valid if the radius a of the DWA is fixed.



(a) (b)

FIGURE 18. (a) Plot of the transformer ratio vs. N (Eqn.(22); (b) plots of dependences of wake potential and the factor I_0 upon N (Eqn.(20) and (21)), with conditions of Q_{total} and R/Q are both held constant.

As shown in Fig.17, the triangular bunch has a broader bandwidth than a Gausiian bunch so that it is difficult to operate in single mode. However, we can design the structure in such a way that R/Q of the HOM's is smaller, thus decreasing their effect on the transformer ratio TR. Table IV shows major parameters of the DWA structure and Table V shows major parameters of the beam used in Fig.7 at the room temperature.

Table IV. Major parameters of the DWA structure used in Fig. 7 layout

ID, OD, Ls	400μm, 464.7μm, 10cm
ε_r , $tan\delta$	3.75, 0.6x10^{-4}
Freq. of TM_{01}, TM_{02}, TM_{03}	850GHz, 3092GHz, 5749GHz, 8489GHz
Q of TM_{01}, TM_{02}, TM_{03}	1260, 3173,4401,5304
r/Q of TM_{01}, TM_{02}, TM_{03}	94.1kΩ/m, 3.2kΩ/m, 0.5kΩ/m, 0.16kΩ/m
Vg of TM_{01}, TM_{02}, TM_{03}	0.592c, 0.794c,0.813c, 0.817c
Q_{total}	1.6nC/ drive bunch
Drive bunch length (total), T	3.3ps (1mm)
Unloaded Gradient	114MV/m
Transformer ratio	16.5

Table V. Major parameters of the beam.

Q_{total}	1.6nC/ drive bunch
Drive bunch profile	Double Trianglular bunch
Drive bunch length (total), T	3.3ps (1mm)1
Unloaded Gradient	114MV/m
Transformer ratio	16.5

The wakefield driven by the double triangular drive bunch is shown in Fig. 19. The black dotted line is the wakefield due to the first mode only and the blue curve includes the first 4 modes. The attenuation and group velocity effect for each mode are icluded [20]. Fig. 19 shows that there are a few particular delay times where the wakefield gradient is larger than the single mode case. Also, note that the self wake inside the double triangular bunch does not appear to significantly change, which means the transformer ratio remains very similar.

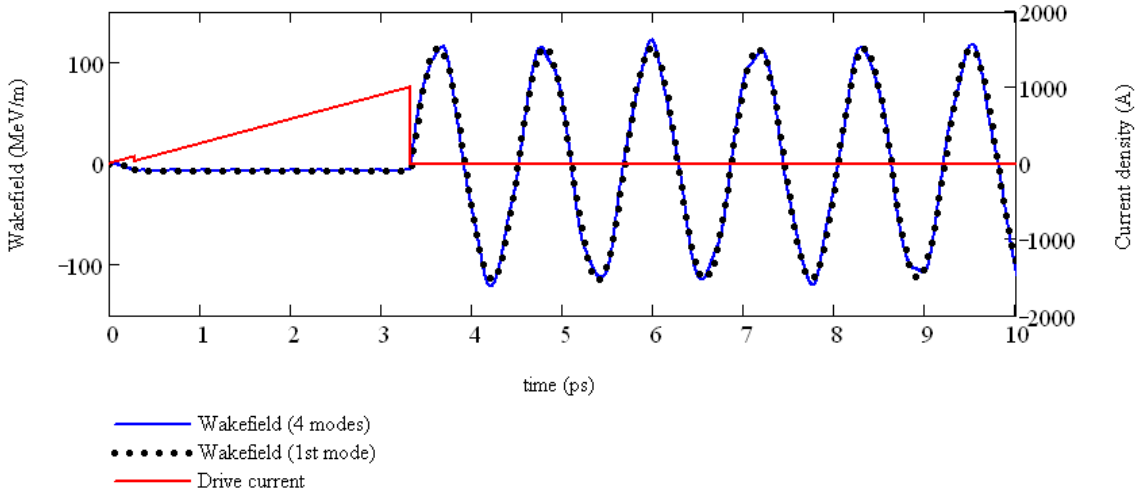


FIGURE 19. Comparison of the wake potential per unit length due to the drive bunch using the fundamental mode only vs. the first 4 modes. There is no need to include more modes because the loss factor of the modes beyond the 4th one is insignificant. Bunch and structure parameters are taken from Fig.7.

VI. Beamloading and Energy Spread

In a collinear wakefield acceleration scheme, the structure radius, *a*, is chosen to be small so that the drive beam excites a strong wakefield. However, because the bunch length of the main bunch is shorter than that of the drive, its self wake (deceleration) can be severe so single bunch beam loading must be considered. Figure 20 shows the overall wakefield plot with a main bunch trailing behind the drive bunch. Due to the main beam short bunch length (~5μm, r.m.s), we included 10 modes in the calculation even though only the first 5 modes contribute to the main beam wakefield.

One way to reduce the beam loading would be to increase the bunch length of the main beam. One could imagine accelerating a long bunch length main beam through the accelerator and then compressing it before the entrance to the FEL. However, another factor to consider is to ensure the energy spread does not get too large for the FEL application. A question that needs to be studied is to what extent this effect can be offset by running off crest through the main linac and compressing the beam before the FEL.

Given the charge distribution of the main bunch, $\lambda(z)$, the energy loss of the centroid (i.e. mathematical expectation) of the main beam, μ_E, is obtained by

$$\mu_E = \int_{-\infty}^{\infty} W(z)\lambda(z)dz\,, \tag{23}$$

where $W(z)$ is the wake potential of the drive bunch. The induced energy spread of a monoenergetic injected beam (i.e. standard deviation) is

$$\sigma_E = \int_{-\infty}^{\infty} \left[\left(W(z) - \mu_E \right)^2 \lambda(z) \right]^{\frac{1}{2}} dz \,. \tag{24}$$

Using the data shown in Table 4, ~10 MeV energy gain per 10-cm long structure is obtained and energy spread induced is ~150 keV, ~1.5%.

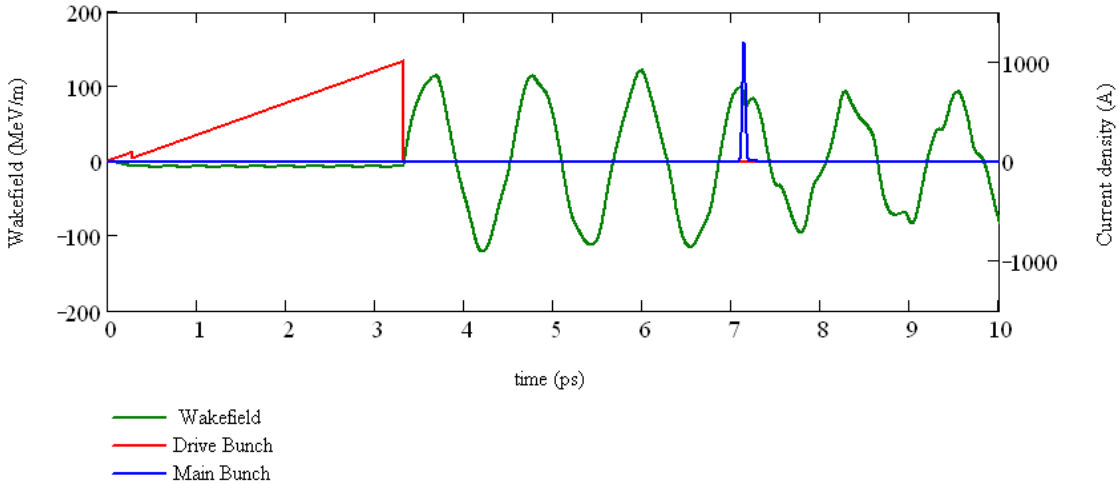


FIGURE 20. Wakefield due to drive and main bunches-.

VII. Thermal Estimation and Cooling

The final subject studied in our straw man design was structure heating. Note that in the ~THz region, the rf attenuation is very high so that the wakefield power will be exhausted in 10 cm (91% of power is dissipated in the structure). This means that the rf packet generated by a drive bunch is only 333 ps long. The single layer DLA structure has a low electric field in the dielectric (E_s ~$1.7E_a$, where E_a is the accelerating field) but a high magnetic field on the metal surface [11]. Therefore, the rf power is mostly dissipated in the copper surface and dielectric losses contribute little for the low loss materials considered here. For example, the quality factor Q of the TM_{01} mode of the structure presented in Fig.7 is 1260/1300 with/without taking dielectric losses into account.

Two types of thermal issues were considered when designing the accelerator: rf pulsed heating (mostly for the pulsed rf machine) and average thermal heating. Pulse heating is an instantaneous phenomenon occurring primarily within a skin depth of the metallic surface. The main concern with pulsed heating is that it might lead to rf breakdown due to contraction and expansion of the metal surface caused by pulsed heating which causes the metal to fatigue. Rf pulse heating usually is not an issue for wakefield accelerators due to the short rf pulse length from the drive bunch. Following the procedure in [38], the temperature rise from rf pulse heating is only 17°C (conditions: 100 MV/m of the loaded gradient→682 kA/m H_{max} at the copper surface, 100 ps square pulse length (to simulate the 333 ps heavily attenuated pulse).

Average thermal heating is a temperature rise in the volume of the accelerating structure due to rf power dissipation. It can be ameliorated by the cooling system. For a high repetition rate machine, the average thermal heating will be one of the most important factors to address in the design process. Let us consider a simplified model. The heat transfer function is

$$\vec{\nabla}^2 T - \frac{1}{\alpha_T}\frac{\partial T}{\partial t} = -\frac{1}{k} q_{gen}, \tag{25}$$

where T is the temperature, α_T is the thermal diffusivity (units of m^2/sec), k is the metal thermal conductivity (units of W/(m °K)), and q_{gen} is the generated heat flux. For the DWA structure shown in Fig. 21, and in the equilibrium state, Eqn. (25) can be reduced to the ordinary differential equation:

$$\frac{dT^2}{d^2 r} + \frac{1}{r}\frac{dT}{dr} = -\frac{1}{k} q_{gen}. \tag{26}$$

Its solution is

$$T = -\frac{q_{gen}}{4k} r^2 + A \ln r + B, \tag{27}$$

where A and B are constants which are determined by the boundary conditions. We assume that the wakefield signal in DWA structures is a traveling wave, which means that there is no rf power reflected back from the downstream end of the structure. This is generally valid because the high attenuation of the wakefield signal (refer to the first paragraph of this Section). However, if the quality factor Q improves (e.g. for very low temperature cooling), the wakefield will form a standing wave so that the rf power behind the drive and main beam will eventually die out inside the DWA structures. Adding an output coupler to the end of each DWA structure for the high Q case will help reduce the cooling load but the cost of single DWA structure will be increased.

The rf power of the wakefield signal can be described as [39]

$$P_{rf}(z) = P_0 e^{-2\alpha Z}, \tag{28}$$

where P_0 is the peak power (power calculated based on the voltage at the wake front), α is the attenuation coefficient (unit of $1/m$), and Z represents the distance from the wake front. Therefore, the dissipated rf power per unit length is

$$q_{gen} = \frac{dP_{rf}}{\pi r^2 dz} = -\frac{2\alpha P_{rf}}{\pi r^2}. \tag{29}$$

Combining Eqn. (27) and (29), we can estimate the temperature rise of the DWA structure due to the average rf power dissipation.

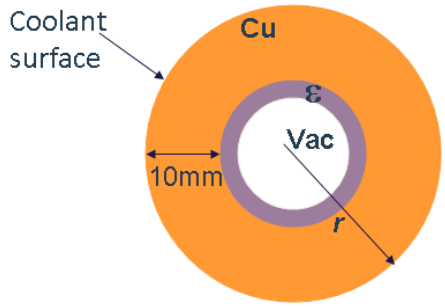


FIGURE 21. Model of the DWA structure used in the thermal estimation. Cooling takes place 1 cm away from OD of the dielectric tube.

As shown in Figure 7, the average dissipated rf power is 5 W/cm^2 (or 7.4 W per structure) at room temperature. This is a relatively high number for this fine capillary structure, but still manageable for the water cooling technology. Figure 22 shows the results of the simple heat transfer model. ~50°K temperature rise above the room temperature occurs at the dielectric surface.

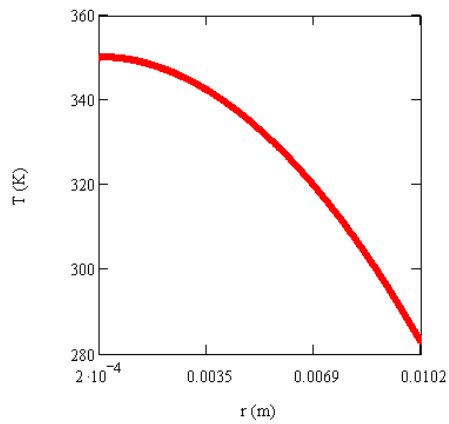


FIGURE 22. Estimated temperature rise from the dielectric surface to the 283°K cooling surface.

Cooling technology is very sophisticated. The thermal estimation model used here was very rough, but one message is clear: sufficient water cooling has to be seriously considered in the design of a high repetition rate wakefield accelerator.

In addition, it is worth considering approaches to reduce the wakefield power radiation without sacrificing the other major parameters (gradient, repetition rate, etc.) One thought is to change the drive beam structure: adding a large charge secondary witness bunch after the main bunch in such a way that wakefield from the secondary witness bunch can mostly cancel the wake thereafter. In other words, the secondary witness bunch absorbs a considerable portion of the rf power left by the drive bunch. From the energy point of view, instead of transferring to the heat load, the leftover wakefield energy from the drive beam (after the acceleration of the main beam) transfer to the kinetic energy of the secondary witness beam. Of course, the efficiency will drop (due to the drive beam current is increased) if the secondary witness beam is eventually dumped. But this could be minimized if its energy were recovered as in an ERL.

VIII. Future

The collinear wakefield accelerator is a potential candidate for providing a beam for the next generation FEL based light source. The requirements for high repetition rate, high gradient, and high efficiency make the design of the accelerator a challenging job, even at the conceptual level. The straw man design proposed here is meant to shed some light on the future work. Several issues were discussed, including transformer ratio, wakefield gradient, bunch shaping, beamloading, thermal heating, etc. More topics have not been touched yet, like beam generation, transportation, etc. Numerous parameters are involved. A good optimization procedure should be set up after the initial nudge of this topic.

References:

[1]. https://slacportal.slac.stanford.edu/sites/lcls_public/Pages/Default.aspx

[2]. http://flash.desy.de/

[3]. http://www.elettra.trieste.it/FERMI/

[4]. http://www-xfel.spring8.or.jp/

[5]. http://xfel.desy.de/

[6]. http://www.sparx-fel.it/index.php/en

[7]. I. Blumenfeld, *etc*, *Nature* 445 Feb. 2007 pp.741-744.

[8]. W. P. Leemans, *etc. Nature Phy.* 2 Oct. 2006, pp.696-699.

[9]. M. C. Thompson, *etc. Phy. Rev. Lett.* 100, 214801 (2008).

[10]. http://clic-study.web.cern.ch/clic-study/

[11]. W. Gai and C. Jing, "The Dielectric-Loaded Accelerating Structures", Periodic Structures, 2006: *ISBN: 81-308-0032-2*, Editors: Maurizio Bozzi and Luca Perregrini.

[12]. Liling Xiao, *et al*, *Phys. Rev. E*, 65, (2001) :016505.

[13]. C. Jing, *et al*, *Nucl. Instr. Meth. in Phys. Res. Sec. A*, 539, No.3, (2005): 445-454.

[14]. G. V. Sotnikov, *et al*, *Proc. 13th AAC Workshop*, Santa Cruz, July 2008, AIP Conf. Proc. **1086**, p. 415(2009).

[15]. K.-Y. Ng, *Phys. Rev. D* (1990):1891-1828.

[16]. M. Rosing and W. Gai, *Phys. Rev. D* (1990):1829-1834.

[17]. H. Braun, CERN-CLIC-Note 364, (1998):150-158.

[18]. D. Whittum, SLAC-PUB-7802, (1998):52-54.

[19]. W. Panofsky and W. Wenzel. *Rev. Sci. Instr.* 27 (1956): 967.

[20]. C. Jing, WF Note, WF-222, http://www.hep.anl.gov/awa/links/wfnotes.htm

[21]. P.B. Wilson, *Proc. the 13th SLAC Summer Inst. On Particle Physics*, SLAC Report No. 296, pp. 273-295.

[22]. J.G. Power, *et al*, *Phys. Rev. E*, **60**, (1999) 6061.

[23]. J. Power, *et al*, WF Note, WF-187, http://www.hep.anl.gov/awa/links/wfnotes.htm

[24]. J. Power, *et al*, *Phys. Rev. E* 60, (1999): 6061-6067.

[25]. Bocheng Jiang, *et al*. a paper is in preparation. Also contact AWA group for details (http://hep.anl.gov/awa).

[26]. W. Gai, A. Kanareykin, A. Kustov, J. Simpson, *Phys. Rev. E* **55**, pp. 3481 (1997).

[27]. C. Jing, *et al*, *Phys. Rev. Lett*., **98**, 144801, 2007.

[28]. C. Jing, *et al*, *Phys. Rev. ST Accel. Beams*., 14, 021302 (2011).

[29]. R. J. England, J. B. Rosenzweig and G. Travish, Phys. Rev. Lett., **100**, 214802, (2008).

[30]. M. Cornacchia, *et al*, *Phys. Rev. ST Accel. Beams* 9, 120701 (2006).

[31]. P. Piot, *et al*, *AIP Conf. Proc.*, 1086, 677-682.

[32]. P. Piot, *et al*, *Phys. Rev. ST Accel. Beams*, *to appear*, *preprint:*ArXiv:1007:4499(2010).

[33]. P. Emma, *et al*, *Phys. Rev. ST Accel. Beams* 9, 100702 (2006).

[34]. Y.-E Sun, *et al. Proc. PAC07*, New Mexico, (2007): 3441-3443.

[35]. K. J. Kim and A. Sessler. *AIP Conf. Proc*. 821, (2006):115.

[36]. www.hep.anl.gov/awa

[37]. K. L. Bane, P. Chen, and P. B. Wilson, SLAC-PUB-3662, 1985.

[38]. C. Jing, WF Note, WF-240

[39]. T. Wrangler, RF Linear Accelerators (John Wiley & Sons, Inc., 1998).