

Self-seeded X-ray Free-Electron Lasers

Gianluca Geloni

Abstract Self-seeding is an active filtering method for Free-Electron Lasers (FELs) enabling the production of nearly Fourier-limited pulses in the X-ray frequency range where external seeding is not available. Schematically, it is composed by three parts: a Self-Amplified Spontaneous Emission (SASE) FEL working in the linear regime, a monochromator and an FEL amplifier. Active filtering is achieved by letting the FEL pulse produced in the SASE FEL through the monochromator, while the electron beam is sent through a bypass and its microbunching is destroyed due to dispersion. The filtered SASE pulse, serving as a seed, is recombined with the electron beam at the entrance of the FEL amplifier part. It is then amplified up to saturation and possibly beyond via post-saturation tapering. This allows for the production of high-brightness, nearly single-mode FEL pulses. The technique has been or will be implemented in a number of X-ray FEL (XFEL) facilities under operation or in the construction phase. In this chapter we review the principle of self-seeding, its practical realizations and related techniques.

1 Introduction, Motivations and Historical perspective

An FEL system acts as an amplifier for a given input signal at a certain frequency. This input signal is completely defined once the electric field, the electron current density modulation and the electron energy modulation at the entrance of the FEL undulator are specified. From a mathematical viewpoint, these three quantities constitute the initial conditions for the FEL dynamical equations and can differ very much depending on the FEL system configuration (see e.g. the textbook [1]). For an externally seeded FEL, an electric field can be provided e.g. by injecting a laser pulse into the system, while the initial electron current density and energy modula-

Gianluca Geloni
European XFEL GmbH, Albert-Einstein-Ring 19, 22761 Hamburg, Germany, e-mail: gianluca.geloni@xfel.eu

tion can often be neglected. In this case, the FEL works as an amplifier for a fully transversely and longitudinally coherent input signal.

In contrast to that, in a SASE FEL [2, 3], it is the current density modulation arising from shot noise that serves as an input signal, while the initial energy modulation can be neglected and there is no initial input field. Therefore, in a SASE FEL the input signal is a realization of a Gaussian stochastic process, and it is constituted by many transverse and longitudinal modes.

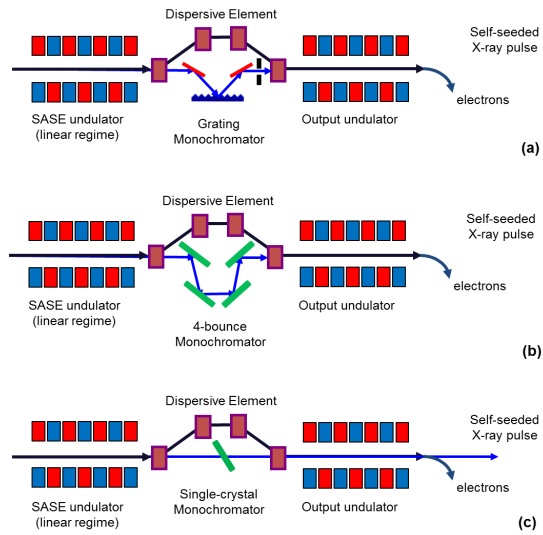
The FEL amplification process itself is parametric. This means that once the undulator parameter is specified, amplification still depends on the electron beam properties: energy, current, energy spread and emittance as a function of the position inside the electron beam. Only once these properties are fixed the response of the FEL to an input signal at a given frequency is fixed as well, and can be described in terms of a superposition of well-defined self-reproducing transverse modes propagating along the undulator with different gains. Because of this difference in gain, if the undulator is long enough one mode dominates over all the others and the output radiation from a SASE FEL is nearly fully transversely coherent, even though the input signal is not.

In contrast, no mode selection takes place in the longitudinal direction. As a consequence, if the FEL process begins with a coherent input signal the output is characterized by an outstanding longitudinal coherence too, but if it starts from an incoherent input, like the noise density modulation in the SASE case, the output also consists of many independent longitudinal modes. Note that the resonance condition imposes that radiation slips forward with respect to the electron beam, one wavelength every undulator period. Due to the exponential nature of the amplification process, this fact roughly limits the build-up of any longitudinal correlation of the electric field during the FEL process to the scale of a number of wavelengths equal to the number of undulator periods within a gain length. In other words, different parts of the electron beam lase in a completely independent fashion.

Currently, the limits to the longitudinal coherence in short-wavelength FELs operating below the nanometer range cannot be overcome with the help of an external seed due to the lack of coherent X-ray sources powerful enough to compete with the equivalent shot-noise level of the electron beam. Because of this reason, up to recently, only rather incoherent SASE FEL radiation was available to experiments at those short wavelengths. Yet, many experimental applications benefit from radiation with a high degree of longitudinal coherence. Self-seeding answers the need for improving longitudinal coherence of FELs in the short-wavelength range mentioned above, to which we will limit our considerations in this chapter.

Of course, the spectral bandwidth of SASE FELs can be narrowed with the help of crystal monochromators, which are installed as standard equipment in many instruments. However, the use of monochromators just corresponds to passively filtering the output radiation and results in a decrease of the pulse spectral density. In contrast to that, self-seeding can be considered as a kind of active filtering process, where the SASE radiation is first passively filtered, and then amplified again. In this way, it is the radiation from the FEL itself which, once filtered, acts as a seed, hence the denomination ‘self-seeding’.

Fig. 1 Possible realizations of a self-seeded FEL. (a) grating monochromator, (b) four-crystal monochromator, (c) single crystal or ‘wake’ monochromator.



The original self-seeding idea can be found in [4]. In general, a self-seeded FEL consists of three parts. The first part is an FEL working in SASE mode and in the linear regime. In the second part the radiation pulse and the electron beam are separated and manipulated in such a way that the electron beam density modulation built up during the previous FEL process is destroyed with the help of a dispersive element, while the bandwidth of the radiation pulse is narrowed down, so that it can serve as a seed. Finally, the third part is an undulator working as an FEL amplifier. Possible realizations of self-seeding differ, depending on the way radiation is monochromatized in the second part of the setup. Fig. 1 shows the case of a grating monochromator (a), of a four-crystal monochromator (b) and of a single crystal or ‘wake’ monochromator (c). Solution (a) (with a modified design compared to that shown in the picture) and (c) were built in practice and will be discussed in detail in the following sections.

Note that, independently of the setup design, the seed should be strong enough to overcome the equivalent shot noise power in the electron beam, and the electron beam quality at the entrance of the last part of the self-seeding setup should be good enough to allow for efficient amplification to take place. Then, in the third part of the self-seeding setup the seed is superimposed with the electron beam, which is still a good active medium, and is thus amplified. In this way, the spectral brightness of a conventional SASE FEL is increased manifold. More precisely, if the peak power of a self-seeded pulse is the same as that of a SASE pulse, the spectral brightness of the self-seeded pulse is increased of a factor roughly equal to the number of modes in the SASE pulse.

The feasibility of the scheme heavily relies on the extremely good quality of the electron beam delivered by state-of-the-art accelerators like those at the largest XFEL facilities under construction or in operation [5, 6, 7, 8, 9]. The length of the

undulators in many XFELs actually exceeds that needed to strictly reach saturation at nominal design energies. This extra available length opens the possibility to implement schemes beyond baseline, like self-seeding.

At the time when [4] was written, no hard X-ray FEL was available. Therefore the self-seeding method was first proposed for the first soft X-ray FEL facility in the world, the Tesla Test Facility TTF, that was under construction at DESY and that now, after upgrades, is known as FLASH [10]. There it was suggested to use a grating monochromator to filter the SASE pulse in the linear regime from the first part of the self-seeding setup, Fig. 1(a). One of the main issues with the realization of that scheme was the limitation of the delay of the radiation pulse relative to the electron beam introduced by the monochromator. This delay must be compensated by the dispersive element, a chicane or a bypass, which delays the electrons and - as discussed above - has also the function of destroying the microbunching accumulated in the first part of the self-seeding setup. If the delay is in the order of tens of picoseconds, a short chicane cannot compensate it, and one needs to resort to a longer bypass [11]. Therefore, great care has to be put in the design of a suitable monochromator. In the soft X-ray range, a grating monochromator specially designed to keep the delay well below a picosecond recently allowed the application of self-seeding in the soft X-ray range around 1 keV at the LCLS [12] using a very compact setup. In the hard X-ray range above 3 keV, crystal monochromators can be used. As suggested in [13], instead of a grating monochromator one may use a four-bounce crystal monochromator, Fig. 1(b). However, also in this case special care must be taken in the monochromator design, in order to limit the delay of the seed with respect to the electron beam. Moreover, the use of conventional monochromators is complicated by the task of aligning the electron beam to obtain proper overlap with the radiation pulse at the entrance of the radiator.

In the hard X-ray range a simpler scheme is available, which is based on a single crystal monochromator [14], Fig. 1(c). This scheme relies on a single diamond crystal, which transmits the spectrum of the impinging XFEL beam almost unvaried, except for a very narrow spectral bandwidth. In other words, the XFEL spectrum transmitted through the crystal presents a narrow dip. In the time domain, the presence of such dip corresponds to a monochromatic ‘wake’ following the main XFEL pulse, which travels through the crystal almost unvaried, hence the name ‘wake monochromator’. The wake is intense enough to be used as a seed and follows the main radiation pulse with delays typically smaller than 100 fs. The great advantage of a Hard X-Ray Self-Seeding (HXRSS) setup based on a wake monochromator is its compactness and the simplicity of its alignment. Since the delay to be compensated is smaller than 100 fs, a relatively weak magnetic chicane can be used as dispersive element, and can easily be fitted into a few-meter long setup. Moreover, the transverse alignment of the radiation pulse is automatically granted. Such a setup has been built at the LCLS [15], where a self-seeded mode of operation is available to users, and SACLA [16], and has been planned at all largest XFEL facilities under construction [17, 18, 19].

Here we will give a review of self-seeding techniques at FELs with particular attention, in section 2, to the wake monochromator case. The case of a self-seeding

setup based on a grating monochromator will be considered in detail in section 3. In section 4 we will discuss issues like the sensitivity of the technique on the electron beam quality and the crystal heat-load. In section 5 we will consider the combination of self-seeding with post-saturation tapering techniques. Tapering consists of a change in the magnetic field peak value along the undulator, in order to compensate for the loss of energy of electrons, which goes into radiation [20]-[32]. This technique is particularly advantageous when coupled with self-seeding and will hopefully provide key for increasing the brightness of XFEL pulses of several orders of magnitude [27]-[32]. Finally, we will come to conclusions.

2 Self-seeding XFELs with a wake monochromator

As already remarked in the Introduction, the original self-seeding concept [4, 13] based on conventional monochromators, Fig. 1(a) and (b), is complicated by the fact that a very long bypass is needed in order to compensate the delay of the electron beam¹ introduced by the monochromator [11]. In fact, the bypass not only needs to compensate the large delay, but must also be nearly dispersion-free, with matched beta functions at the entrance and at the exit, and with the possibility of fine tuning the extra path length. If no special design solutions are adopted, studies for e.g. the FLASH case showed that several tens of meters are required, a large modification to the baseline of any facility [11].

A possible way to overcome this issue was suggested [33, 34]. It amounts to seeding an electron bunch with radiation generated by a previous one, based on a photoinjector setup with a laser pulse doubler. Only a short chicane is then needed. A first experiment for integrating the laser pulse doubler with an XFEL photoinjector was performed at FLASH and reported in [35]. A pulse doubler was recently installed at the LCLS, allowing acceleration of two bunches in the same RF bucket [36].

However, a simpler solution is available for hard X-ray energies above 3 keV, based on a special monochromator known as ‘wake’ monochromator [14], Fig. 1(c). With this kind of monochromator one can keep the delay around 100 fs and below (depending on the electron bunch duration, as we will see). This allows for building a self-seeding setup with a length limited to a few meters only. The use of a wake monochromator also dramatically simplifies the tuning of the entire setup. Such choice is currently the state-of-the art for self-seeding in the hard X-ray range. The reasons for the limitation in energy is due to the crystal-based technology upon which the practical realization of the wake monochromator relies.

We will first discuss the principle of operation of a wake monochromator from a very general standpoint, without considering, as in Fig. 1(c), its particular realization based on a single diamond crystal. In fact, we find that the approach of decoupling

¹ In contrast to the delay issue, for an energy spread of order $\Delta\gamma/\gamma \sim 0.01\%$ and wavelengths in the order of 0.1 nm, the electron beam microbunching is washed out already for an R_{56} in the order of a few microns.

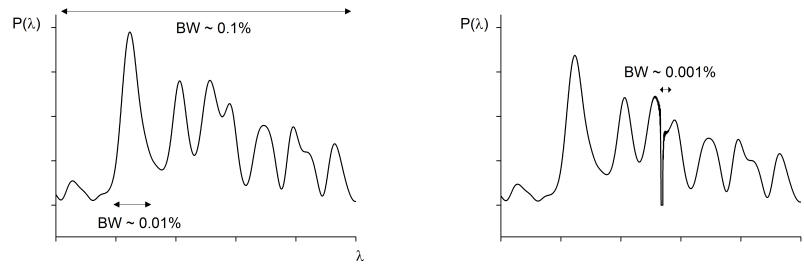
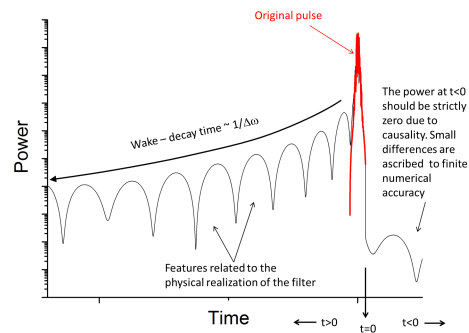


Fig. 2 Example of radiation spectra before (left) and after (right) a wake monochromator. The filtering process, aside for absorption effects, mainly amounts to drilling a hole in the spectrum, exactly at the seeding frequency.

Fig. 3 Example of transmitted power in the time domain. The red pulse is the original incident pulse. The behavior of the trailing 'wake' is clearly visible in logarithmic scale.



the general idea from the practical realization is more interesting from a theoretical viewpoint.

2.1 Principles of Operation

The wake monochromator principle is based on the idea of placing a bandstop filter on the path of the FEL pulse from the first part of the self-seeding setup so that, after transmission, the spectrum exhibits a narrow dip in the bandstop region. This dip should be narrower than the average spike in the FEL spectrum and must be tuned exactly at the frequency where one wants to seed, as illustrated in Fig. 2. This last fact is obviously paradoxical, and is better explained in the time domain. Let aside for a moment the dependence on the transverse coordinates, the relation between frequency and time domain can be qualitatively explained in the following way. In the frequency domain the field of the transmitted X-ray pulse is given by the product of the incident field and the transmissivity of the filter, which are both complex functions of the frequency with given modulus and phase. In the time domain, instead, the transmitted X-ray pulse can be seen as a convolution between the

incident X-ray pulse and the inverse Fourier transform of the filter transmissivity, which behaves as impulse response for the entire system, and will be called \mathcal{T} .

Schematically, in the frequency domain a bandstop filter can be represented as the difference between two terms. The first is an almost constant term, ideally equal to unity if one neglects absorption related to the filtering process, and models the fact that the signal is transmitted without variation outside of the bandstop region. The second term consists of a narrow window function of the frequency multiplied by a phase factor. The exact shape of this second term depends on the physical realization of the filter. By linearity of the Fourier transform, one can analyze the two terms separately in the time domain and then add them together. The Fourier transform of the first term yields a Dirac- δ function. Convoluting it with the incident signal gives back the original signal in the time domain. This is intuitively sound: if there is no bandstop filter, one obtains the original signal. The Fourier transform of the second term yields, instead, a function which decays to zero with a characteristic time of order of the inverse of the bandwidth of the bandstop filter. The phase of the second term is responsible for causality, meaning that if the original signal starts at $t = 0$, the convolution with the original signal must be zero when $t < 0$. The role of the phase is also intuitive, if one remembers that the derivative of a frequency phase gives a group delay at a given frequency. As a result, one obtains an impulse response that is composed of a sharp, Dirac- δ function followed by a tail decaying with the inverse of the bandwidth of the bandstop filter.

The field transmitted in the time domain is the convolution product of the impulse response with the initial electric field pulse in the time domain. From our previous discussion it follows that such a field consists of a main radiation pulse reproducing the initial one followed by a tail, Fig. 3. The monochromatization process is completed by superimposing the electron bunch on top of this tail. This operation is mathematically equivalent to a multiplication by a temporal window of duration equal to the bunch duration, centered on the radiation tail. Note that without this last operation there would be no monochromatization process at all. An alternative view of this process is a spectral measurement performed with a very short measuring time, the bunch duration, which limits the spectral range of the measurement, actually enforcing monochromatization.

Let us now write the transmissivity as $T(\omega) = |T(\omega)| \exp[i\Phi(\omega)]$. While in the case of a crystal-based realization (see section 2.2) one can directly determine $T(\omega)$ using the dynamical theory of X-ray diffraction, from a theoretical viewpoint it is interesting to consider the interdependence between $|T(\omega)|$ and $\Phi(\omega)$. As discussed above, the impulse response of the system, which we called $\mathcal{T}(t)$, must grant causality of the transmitted pulse. The assumption of square-integrability of \mathcal{T} , has far-reaching consequences. One can write

$$\ln[T(\omega)] = \ln[|T(\omega)|] + i\Phi(\omega) . \quad (1)$$

Assuming $\mathcal{T}(t)$ real one has $T^*(\omega) = T(-\omega)$ so that $|T(\omega)| = |T(-\omega)|$ and $\Phi(\omega) = -\Phi(-\omega)$. Therefore, using Eq. (1) it can be seen that $\ln[T(\omega)]^* =$

$\ln[T(-\omega)]$. We can now invoke Titchmarsh theorem (see [37] and references therein for a review) to ensure the existence of an analytic continuation of $\ln[T(\omega)]$ to $\Omega = \omega + i\xi$ for $\xi > 0$ as well as the fact that a Hilbert transform relates $\text{Re}[\ln(T)]$ and $\text{Im}[\ln(T)]$ in analogy with Kramers-Kronig relations [38, 39]. From the Hilbert transform one obtains the following relation between the phase and the modulus of the transmission function [40]:

$$\Phi(\omega) = -\frac{2\omega}{\pi} \mathcal{P} \int_0^\infty \frac{\ln[|T(\omega')|]}{\omega'^2 - \omega^2} d\omega' . \quad (2)$$

In other words, Eq. (2) allows direct calculation of the phase of the transmission function, $\Phi(\omega)$, once its modulus $|T|$ is known. It should be noted, however, that in applying such a procedure one tacitly assumes that $\ln[|T(\Omega)|]$ is analytical on the upper complex Ω -plane. Causality implies this for $T(\Omega)$, but not for $\ln[|T(\Omega)|]$, which is a singular function whenever $T(\Omega) = 0$: possible zeros on the upper complex plane contribute with extra terms to the phase. This is why Eq. (2) is known as *minimal* phase solution.

However, the passage through a filter with transmission function $T(\omega)$ can be very often modeled as the passage through a slab of matter with refractive index n and thickness d according to

$$T(\omega) = \exp \left[\frac{i\omega}{c} n(\omega) d \right] , \quad (3)$$

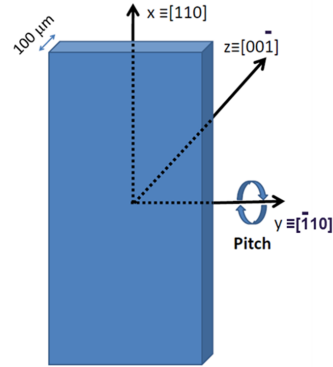
implying

$$\ln[T(\omega)] = -\frac{\omega d}{c} \text{Im}[n(\omega)] + i \frac{\omega d}{c} \text{Re}[n(\omega)] . \quad (4)$$

Since the Kramers-Kronig relations automatically hold for the real and the imaginary parts of the refractive index (see e.g. [37], Paragraph 4.6.1 for a textbook explanation), the validity of Eq. (2), and hence the possibility of calculating the phase of the transmissivity upon the knowledge of its modulus, follows under the quite general assumption of validity of Eq. (3), independently of the particular physical realization of the filter. An application of this fact is shown in the Appendix, where we consider an exotic realization of a wake monochromator for the long wavelength range based on a ‘gas-cell’ monochromator. Due to several limitations also discussed in the Appendix this scheme remains, up to now, only of theoretical interest.

To close this section, we should consider that the discussion made up to now assumes no dependence of field and transmissivity on transverse coordinates. A more general treatment is possible, where one should consider the transmissivity as a function of the temporal frequency and of the momentum vector of each plane-wave component of the incident radiation, separately filter each plane-wave component of the incident field and, finally, synthesize the outcoming field. However, the depen-

Fig. 4 Drawing of the orientation of the diamond crystal used at the LCLS.



dence of the incident field on transverse coordinates can be readily included in the present treatment. Moreover, it should be noted that the incoming XFEL beam is very well collimated, in the order of a microradian, which is usually much narrower than the angular acceptance of the filter.

In contrast to that, the spectral bandwidth of the incoming beam is much wider than the bandwidth of the filter transmissivity, which is a function of the incident angle of the radiation. As we will see in the next section, in the practical realization of the wake monochromator based on a single crystal, this leads to a particular coupling between transverse and longitudinal coordinate in the expression for the filtered field.

2.2 HXRSS with a single-crystal wake monochromator

Above 3 keV, a single crystal plate in transmission geometry can be used in order to obtain one particular realization of the filter described in the previous section. We will designate this energy range as ‘hard X-rays’. The incident beam hits the crystal at the Bragg angle corresponding to the seed frequency for a given reflection. Part of the radiation in a very narrow bandwidth is reflected, while the transmitted spectrum shows the narrow dip which is needed to implement the self-seeding wake monochromator scheme. Due to low absorption, the choice of diamond is optimal in order to minimize heat-loading effects. The dynamical theory of X-ray diffraction is the natural tool to be used in order to compute modulus and phase of the transmissivity of the filter although, as discussed before, modulus and phase are not independent of each other, but are linked by the Kramers-Kronig relations as verified also in [14].

Fig. 5 Scattering geometry of the C(004) Bragg symmetric reflection at 8 keV from the crystal in Fig. 4.

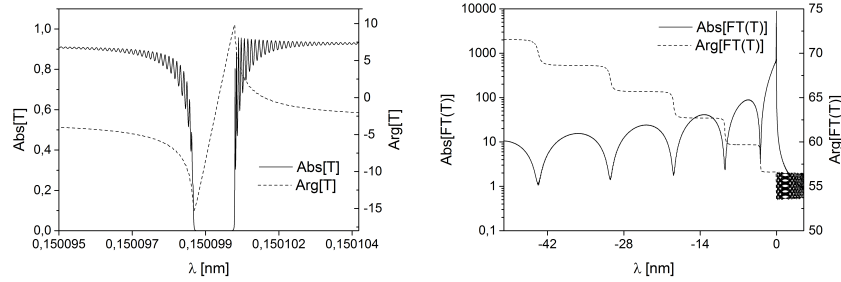
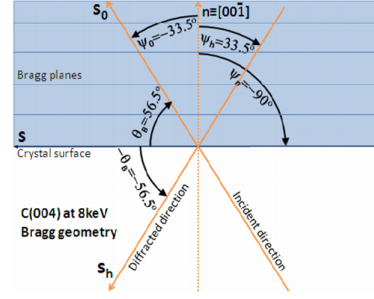


Fig. 6 Transmissivity and impulse response for the C(004) Bragg symmetric reflection at 8 keV from the crystal in Fig. 4. Note that the phase jumps abruptly of 2π radians in the impulse response (right plot).

Depending on how the crystal is cut and on the seed energy, we can exploit both symmetric and asymmetric crystal reflections. This fact was first realized at the LCLS and discussed in detail in [41, 42] after successful operation of the first HXRSS setup, where the $100\ \mu\text{m}$ thick diamond crystal proposed in [14] and shown in Fig. 4 was used [15].

By rotating the crystal around the y axis (pitch axis) one can not only finely tune the seed frequency for a fixed crystal plane, but also exploit several symmetric and asymmetric reflections. This allows one to cover any energy between 3 keV and 13 keV using only five different reflections, see [42] for actual data: the C(111) asymmetric reflection (in Bragg or Laue geometry, depending on the energy) above 3 keV, the C(113) asymmetric reflection in Bragg geometry above 6 keV, the C(004) symmetric Bragg reflection above 7 keV, the C(333) asymmetric reflection in Bragg geometry above 9 keV, and the C(444) symmetric Bragg reflection above² 12 keV. A more detailed treatment of the different reflections is reported in [43]. Here we limit ourselves to show, for illustration purposes, the scattering geometry, the transmissivity and the impulse response pertaining the C(004) Bragg symmetric reflection at 8 keV, Fig. 5 and Fig. 6. As discussed before, and as showed qualitatively in Fig.

² Above 12 keV the efficiency of self-seeding tends to decrease due to a combination of different reasons linked to the Bragg angle, the width of the reflection, and the gain length.

2 and Fig. 3, the electric field after the first FEL part and the filter in Fig. 6 is obtained in the frequency domain by multiplication with the transmissivity in modulus and phase, while in the time domain the field corresponds to a convolution with the impulse response. An overall decrease of the wake signal on a temporal scale of the inverse Darwin width of the transmissivity is expected, but in addition to that both the impulse response and the filtered field show ‘bump-like’ features related to the physical realization of the filter which are explained in detail in terms of dynamical theory of diffraction [42]. From a practical viewpoint, the duration and the amplitude of these features are linked to the crystal thickness, which becomes an important parameter to consider in the optimization of the seed, and not only in relation to absorption issues.

Finally, it is very important to remark that a wake monochromator introduces a particular spatiotemporal coupling of the seeded X-ray pulse, which has been analyzed in the framework of the dynamical theory of X-ray diffraction [42] and is relevant to self-seeding. Due to this distortion, the X-ray beam spot size is independent of the longitudinal coordinate inside the beam, while the transverse position of the beam centroid is not. In [43] it is shown that, based on the Bragg law, one can directly provide an explanation of spatiotemporal coupling phenomena. In fact, since the transmitted and the incident beam have the same direction we can set $\theta_i = \theta_d$ and one obtains the following expression for the transmissivity T :

$$T(\Delta\omega, \Delta\theta_i) = T(\Delta\omega + \omega_B \Delta\theta \cot(\theta_B)) , \quad (5)$$

where $\Delta\omega = (\omega - \omega_B)$ and $\Delta\theta = (\theta_i - \theta_B)$ are the deviations of frequency and incident angle of the incoming beam from Bragg frequency and Bragg angle respectively. The frequency ω_B and the angle θ_B are related by the Bragg law: $\omega_B \sin(\theta_B) = \pi c/d$ and the transmissivity is invariant under angle and frequency transformations obeying $\Delta\omega + \omega_B \Delta\theta \cot(\theta_B) = \text{const}$, also a manifestation of Bragg’s law. This corresponds to a coupling in the Fourier domain. In general, one would expect the transformation to be symmetric in both the Fourier domain $\{k_x, \omega\}$ (with $k_x = \omega_B \Delta\theta/c$) and in the space-time domain $\{x, t\}$, due to the symmetry of the transfer function. However, it is reasonable to expect the influence of a nonsymmetric input beam distribution. The field transformation for the XFEL pulse after the crystal in the $\{x, t\}$ domain is given by

$$E_{\text{out}}(x, t) = FT\{T[\Delta\omega, (k_x)_{\text{in}}, (k_x)_{\text{out}}]E_{\text{in}}(\Delta\omega, k_x)\} , \quad (6)$$

where FT indicates the Fourier transform from the frequency domain $\{k_x, \omega\}$ to the space-time domain $\{x, t\}$. In the self-seeding case, the incoming XFEL beam has an angular spread a few times smaller than the angular width of the transfer function³. However, the incoming beam is characterized by a spectral bandwidth

³ It is indeed this narrow collimation that allows to assume that all frequency components of the incident field impinge at the same constant angle.

which is much wider than the bandwidth of the transfer function. Under this limit and applying the shift theorem twice one gets

$$\begin{aligned} E_{\text{out}}(x, t) &= \eta(t) \int_{-\infty}^{\infty} dk_x \exp(i \cot \theta_B k_x c t) \exp(-ik_x x) E(\omega_B, k_x) \\ &= \eta(t) a(x - c \cot \theta_B t), \end{aligned} \quad (7)$$

where $\eta(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \exp(-i\omega t) T(\omega)$ is the inverse temporal Fourier transform of the transmittance curve (that is the impulse response).

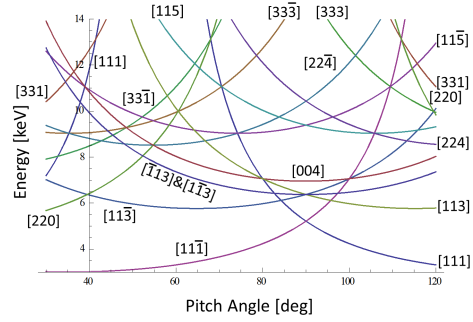
Eq. (7) is universal i.e. valid for both symmetric and asymmetric Bragg and Laue geometries. It describes our spatiotemporal coupling from a mathematical standpoint. Its consequence on the operation of the self-seeding setup was studied by characterizing the input coupling factor between the seed beam and the ground eigenmode of the FEL amplifier in [43]. Qualitatively it was found that the input coupling factor does not go abruptly to zero when the transverse offset becomes of order of the rms transverse size of the photon beam. This is ascribed to the fact that the photon beam intensity does not drop off as a Gaussian function.

2.3 Experimental realization

2.3.1 LCLS

The first HXRSS setup in the world was realized at the LCLS in 2012 [15]. The self-seeded mode of operation is currently offered to LCLS users besides the baseline SASE. The LCLS HXRSS system is based on a 100 μm diamond crystal as sketched in Fig. 4 and described in section 2.2. The diamond monocrystal was grown at the Technological Institute for Super-hard and Novel Carbon Materials (TISNCM, Troitsk, Russia), and is free to rotate around the pitch axis from 45 to 90 degrees (where 90 degrees indicate the FEL beam at normal incidence). The diamond is placed in a vacuum chamber designed in collaboration with Argonne National Laboratory [44]. The crystal is positioned in the 2.5 mm transverse gap created by a 3.2 m-long chicane, capable of creating a variable electron beam delay up to 40 fs, and installed at the sixteenth LCLS undulator module *U16*. A much finer phasing between electrons and radiation is granted by bend trim coils, allowing steps of 0.04 Angstrom, from zero to 10 Angstrom. The setup was first commissioned around 8 keV with the help of a Bragg camera to monitor the C(004) reflection from the crystal as originally suggested in [45], and of a single-shot spectrometer using a Si(333) bent crystal as dispersive element, allowing for a resolution of about 0.1 eV [46]. Reference [15] describes the production of single-shot seeded pulses of about the same energy of SASE (around 240 μJ), but with a FWHM bandwidth of about 0.4 eV, compared to a SASE FWHM bandwidth of about 20 eV. An increase of the average FWHM seeded bandwidth up to about 1 eV was reported and ascribed to shot-to-shot jitter of the electron beam quality. A relative electron

Fig. 7 Many reflections are available by tilting the diamond crystal in Fig. 4 at a certain pitch angle for a given energy. Those with indexes (h_1, h_2, l) and (h_2, h_1, l) with $h_1 \neq h_2$ are superimposed one on top of the other.



energy jitter of about $5 \cdot 10^{-4}$ was found to be responsible for shot-to-shot intensity variations larger than 50%. The LCLS team is working towards a reduction of the electron energy jitter by tracking and correcting its causes. They recently reported a decrease down to a typical relative jitter of $3 \cdot 10^{-4}$ at energies around 8 keV [47]. Brightness was also reported to be increased since the commissioning. In particular, measurements performed with a four-bounce Si(111) monochromator showed that the HXRSS mode of operation offers at least 3 times more average intensity than SASE optimized for maximum pulse energy [48].

Following successful demonstration around 8 keV, operation of the HXRSS setup at the LCLS was extended to other crystal reflections, which allowed wider tunability in energy. Many more reflections than those discussed in section 2.2 are actually available, Fig. 7. It was soon realized [49] that some of these reflections, namely those with indexes (h_1, h_2, l) and (h_2, h_1, l) with $h_1 \neq h_2$ are superimposed one on top of the other (see the $[\bar{1}13]$ and $[1\bar{1}3]$ traces shown in Fig. 7 as an example). However, the additional presence of a yaw rotational stage (between -2.5 and 2.5 degrees at the LCLS) allows one to decouple these reflections, and to control the separation in energy between them at a fixed pitch angle just by tuning the yaw. This opened the possibility of generating seeded pulses of multiple colors at the LCLS [49], which are obtained by seeding simultaneously at energies within the FEL bandwidth. It was shown that the energy separation could be further increased at the LCLS by introducing a strong chirp in the electron beam energy. This method also introduces a temporal delay between the two colors due to the temporal separation of the two parts of the beam with different energy levels. Energy separations larger than a percent can be obtained by accelerating twin bunches in the same RF bucket, [36]. In this case, tunability of the delay between different colors is also granted. An alternative way to produce multiple colors by filtering the initial SASE pulse with multiple crystals tuned at slightly different energies was also theoretically discussed in [50], which allows, as a byproduct, to produce intense bursts of coherent radiation in the optical regime, at an energy equal to the difference between the two colors.

Finally, it should be remarked that once the crystal is retracted, the HXRSS chicane at the LCLS has a use of its own as a diagnostics tool, allowing direct measure-

ments of the average photon pulse length in the linear regime. The method was first suggested in [51]. It consists in having two identical parts of an XFEL undulator preceding and following a short chicane lasing in the linear regime, and recording the average energy over an ensemble of radiation shots (for instance with a gas-monitor detector) as a function of the chicane delay τ . This procedure allows one to recover the autocorrelation function of the ensemble-averaged signal in the time domain $\langle P(t) \rangle$, i.e. $A(\tau) = \int_{-\infty}^{\infty} dt \langle P(t - \tau) \rangle \langle P(t) \rangle$, which in its turn allows for direct measurement of the duration of X-ray pulses of few femtoseconds. Note that this method does not use the HXRSS crystal, and was available from the very beginning of the HXRSS setup commissioning [52]. The self-seeding chicane alone can also be used in order to generate two color SASE pulses (see [53] and references therein).

2.3.2 SACLA

Following the LCLS pioneering demonstration, an HXRSS setup was installed at SACLA between the eighth and the ninth FEL module. The necessary hardware was installed starting from 2012, and the chicane was used for autocorrelation measurements and for the production of 2-color SASE light [54], similarly as discussed in [51] and [55], prior to the installation of the HXRSS crystal in 2013.

Results of the commissioning at a photon energy of 10 keV and at an electron energy of 7.8 GeV are reported in [16]. In the SACLA setup, the diamond crystal is cut as in Fig. 4, but it is 180 μm thick and is free to rotate between 0 and 180 degree around the pitch axis. The C(004) symmetric reflection was used during commissioning, corresponding to a Bragg angle of about 44 degrees. Delays of the two first maxima in the wake were found respectively around 25 fs and 45 fs, in agreement with theoretical calculations. In contrast to the LCLS, the SACLA undulator has a tunable gap. It was therefore possible to set the Bragg angle and then optimize the K parameter by measuring the Bragg-diffracted intensity on a photo-diode attached (together with a CCD camera coupled with a phosphor screen for spatially resolved measurements of the Bragg diffraction signal) on the 2θ -rotation arm of the crystal camber. Following this step, the spectrum of the SASE radiation transmitted through the diamond crystal was analyzed using a single shot-spectrometer in the experimental hutch. Using a Si(660) crystal, a resolution of about 60 meV was obtained, capable of resolving the dip in the transmitted spectrum created by the passage through the crystal. Finally, the gap of the undulator modules downstream of the HXRSS setup was closed and self-seeded spectra were measured with the help of the same single shot spectrometer, this time using the Si(220) reflection to analyze the signal.

Average self-seeded spectra at 10 keV were reported with a spectral bandwidth of 3 eV, about ten times less than SASE, and four times larger than the SASE background. The probability of observing the seeded signal in a single-shot spectrum was found to be 42%, with an intensity fluctuation of 31% rms. The SACLA experiment underlined a very high sensitivity of the seeding process on the RF accelerator set-

tings, and confirmed the importance of a high stability of the electron beam quality for self-seeding purposes.

3 Self-seeding X-ray Free Electron Lasers with a grating monochromator

3.1 Principles of operation and design

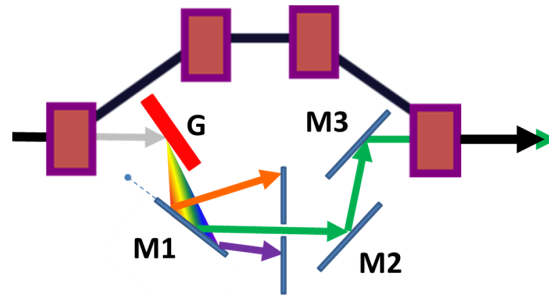
The technical realization of self-seeding in the soft X-ray range moved back to the original idea [4] based on a grating monochromator setup. The setup itself is more complicated than that based on a wake monochromator. Moreover, the use of a conventional monochromator introduces issues in the alignment of the setup. In particular, one needs to provide transverse overlap of electrons and seed at the entrance of the second part of the undulator. In contrast to that, such overlap is automatically granted in the case of a wake monochromator.

However, the use of a grating allows one to overcome the limitations in energy related to crystal-based technology, enabling continuous tunability of the setup around 1 keV. As discussed before, the main challenge with conventional monochromators is the compensation of the delay of the X-ray beam with respect to the electron beam. Previous designs at TTF and FLASH [11] would have introduced tens of picosecond delays, and relied on very long bypass lines (several tens of meters long) in order to compensate it. This kind of design is clearly not easy to adopt, because it radically changes the entire facility.

Considerable effort was invested in theoretical investigation and *R&D* of an SXRSS setup at the LCLS. In a few years, a collaboration between SLAC, Lawrence Berkeley National Laboratory, and the Paul Scherrer Institut led to design [56]-[58] and implementation [12] of a very compact SXRSS grating monochromator introducing only a short delay of the radiation compared to the electrons, of about 0.7 ps.

The monochromator design differs from that in Fig. 1(a), and is sketched in Fig. 8. It is composed of only three mirrors, slits and a toroidal VLS grating. The chicane acts as a bypass for the electron beam, while the incidence angle of the radiation onto the grating is constant, in order to suppress the influence of the movement of the source point in the first SASE undulator on the monochromator performance. A first mirror (M1) can be pivoted around a fixed point, thus tuning the color going through the slits, which is then extracted through the mirrors M2 and M3. The photon beam is then refocused onto the electron beam at the entrance of the second undulator part (not shown in the plot). The seed signal is attenuated of many orders of magnitude compared to the incident one (typical numbers are $1\mu\text{J}$ for the incident SASE pulse and 1 nJ for the seed pulse), but is enough to seed the electron beam.

Fig. 8 Sketch of the SXRSS setup built at the LCLS [12]. The SASE pulse impinges on the grating G , which acts as dispersive element. The X-ray pulse of a chosen frequency is selected by pivoting the mirror $M1$. The selected color goes through the slits and is further refocused on the electron beam with the help of the mirrors $M2$ and $M3$.



3.2 Experimental realization

The overall setup [12] built at the LCLS is capable of seeding an energy range between 500 eV and 1000 eV, with a resolving power between $2 \cdot 10^3$ and $5 \cdot 10^3$. It fits a 3.4 m-long undulator section and was installed at undulator module $U9$. Designers overcame a number of technical issues in order to engineer a compact setup introducing less than a ps delay in the photon beam. Commissioning challenges included, as for the HXRSS case, the right choice of the length of the first SASE sections: too few segments would have resulted in insufficient seed level, too many would have spoiled the electron beam, not allowing for proper lasing in the second section. In addition, a compromise between seed power and damage threshold for the SXRSS optics had to be met. Other peculiar difficulties for SXRSS included the tracking of the radiation pulse through a complicated setup made of five optical elements, determining the positions of X-ray and electron beam notwithstanding the presence of coherent radiation (which proved to be an issue for the proper functioning of the diagnostics) and overlapping electrons and X-rays both in the transverse and longitudinal direction (the delay introduced in the SXRSS setup, albeit short, is still much longer than that in the HXRSS case).

As for the HXRSS case, the quality of the electron beam and the stability of its characteristics constitute limitations to the quality of the seeded beam. In particular, an rms energy jitter of 100% was observed when all shots were considered, while the jitter was reduced to about 50% in the case when only on-energy shots were selected (while SASE fluctuations are about 25%, due to the use of a slotted foil [59, 60, 34], allowing for control of the electron bunch duration.). The shot-to-shot wavelength stability was about 10^{-4} , or 100 meV at about 1 keV, which had the effect of broadening the average seeded bandwidth from 155 meV to 180 meV FWHM.

Even accounting for these detrimental effects, in the case a hutch monochromator is used to filter both self-seeded and SASE pulses, the authors of [12] report an increase in brightness of a factor 2 – 5 compared to optimized SASE. Such in-

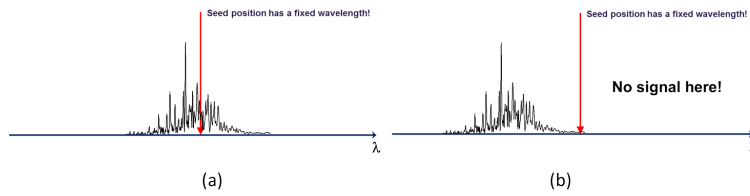


Fig. 9 The seed wavelength is fixed. Depending on the electron beam energy jitter, this wavelength may be within the SASE FEL spectrum, as in plot (a), or outside it, as in plot (b). In this last case, no seed signal is present.

crease reaches a factor 20-50 in the case non-monochromatized self-seeded pulses are compared with monochromatized SASE pulses (as is requested in some experimental applications).

4 Related issues

4.1 Sensitivity on the characteristics of the electron beam

As already discussed, the experiences at LCLS [12, 15] and SACLA [16] underlined the sensitivity of the self-seeded pulses on the electron beam characteristics and their stability, in particular electron energy chirp and electron energy jitter.

Linear energy chirp, that is a linear correlation between electron energy and position inside the electron beam is responsible for a shift of the final seed frequency. Non-linear chirp is responsible for a spread of the spectrum instead. These facts can be understood considering that upon interaction with the undulator field, the seed pulse creates an energy modulation in the electron beam. This energy modulation is then transformed into a density modulation due to the undulator dispersion R_{56} , which induces an energy-dependent trajectory according to $dz = R_{56}d\delta$, where $d\delta$ is the relative deviation from the design energy. It follows that, similarly as in a magnetic bunch compression, a linear energy chirp induces a uniform bunching along the bunch with longer or shorter period compared to design (depending on the slope of the energy correlation) with the overall effect of shifting the central wavelength of the seeded pulse. If, instead, the chirp is non-linear, different parts of the bunch are over- or under-compressed during the bunching process, resulting in a spread of the spectrum. This effect is well-known from the theory of seeded FELs, see [61] and references therein, and was included in HXRSS simulations before the actual experimental verification of the technique, see e.g. [62]. It can be concluded that the presence of non-linear energy chirp can easily constitute the limiting factor to the spectral bandwidth achievable at each shot, which would be otherwise determined by the length of the electron beam according to the Fourier limit.

Electron energy jitter can impact self-seeded pulses inducing, instead, a jitter of the final seeded-pulse energy. In fact, while the crystal transmission is fixed around a certain frequency, the presence of electron energy jitter changes the central frequency of the FEL amplification bandwidth. If the relative jitter becomes larger than about $\rho/2$, where ρ is the FEL efficiency parameter, the central frequency of the crystal transmission does not always superimpose with the FEL spectrum, and not all pulses are seeded. This leads to an increase in the jitter of the final seeded-pulse energy compared to the unavoidable jitter level related to the start-up from shot noise in the first part of the undulator, Fig. 9. This effect was first pointed out by the LCLS staff [15] upon realization of the first self-seeding setup and affects both HXRSS and SXRSS [12] setups.

4.2 Heat loading and high repetition rate

Operation of self-seeding setups described up to now both in the hard and in the soft X-ray regions is limited to low repetition rates in the order of hundred pulses per second, which is what is currently achieved at SASE XFEL facilities like the LCLS and SACLA [5, 6].

However, extension of these techniques to much higher repetition rates is needed in order to match the typical operation modes of XFELs like the European XFEL and LCLS-II [7, 63], which will be driven by superconducting accelerators. For example, the European XFEL will provide up to 27000 pulses of X-rays per second distributed in ten macropulses, each of them consisting of 2700 shots at a frequency of 4.5 MHz [7].

In this case, the heat load on the optical elements can easily become an issue. Here we will consider, as a particular example⁴, the case of a diamond crystal in a wake monochromator. Heat-load effects become unacceptable as soon as the central frequency of the transmission filter is shifted by the expected spectral width of the self-seeded pulse. In fact, in this case one would notice a widening in the output spectral bandwidth averaged over an ensemble of FEL pulses. Usually, a more stringent condition is formulated, requiring that the central frequency of the crystal experiences a shift of less than a Darwin width. Preliminary theoretical studies conducted at the European XFEL [64] showed that a limit to the acceptable energy per pulse around 8 keV (assuming the C(004) symmetric reflection in the diamond crystal sketched in Fig. 5) is about 3 μJ . In particular, it was shown that the drift in the central frequency of the crystal transmission is still within a Darwin width when 1000 pulses with a transverse size of 35 μm FWHM and at a repetition rate of 4.5 MHz, each carrying about 3 μJ energy, impinge on a 100 μm thick diamond crystal. In this case, a cooling system might be needed, but the heat deposited can be removed without major problems.

⁴ For high repetition rates, heat loading studies will also be needed in the case of grating monochromators in the soft X-ray spectral range.

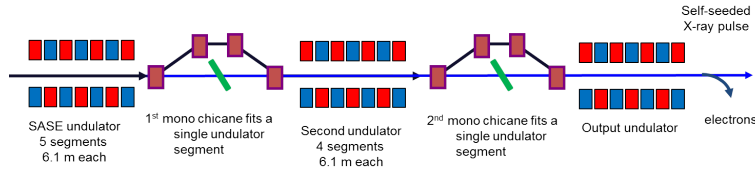


Fig. 10 Possible two-chicane HXRSS setup for the SASE2 undulator line at the European XFEL. The SASE2 line is composed of 35 undulator segments. Each segment hosts a undulator with an effective magnetic length of 5 m and period $\lambda_w = 40$ mm, together with an intersection about one meter long including focusing elements and phase shifters. Each HXRSS setup can be placed in a single segment.

Fig. 11 Incident power and spectrum of the X-ray pulse on the second crystal (C(004) reflection, diamond crystal as in Fig. 5). These plots refer to the two-chicane HXRSS scheme for the SASE2 line at the European XFEL illustrated in Fig. 10. The grey lines represent 100 simulations, with different initial shot-noise conditions. The black line is the average over these 100 shots.

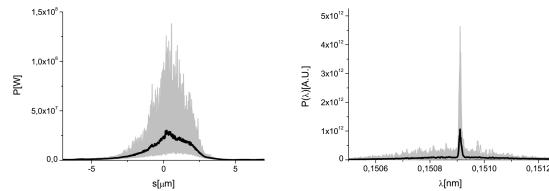
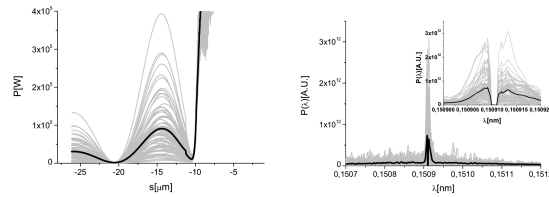


Fig. 12 Seed power and effect of the second crystal (C(004) reflection, diamond crystal as in Fig. 5) on the input spectrum. These plots refer to the two-chicane HXRSS scheme for the SASE2 line at the European XFEL illustrated in Fig. 10. The grey lines represent 100 simulations, with different initial shot-noise conditions. The black line is the average over these 100 shots.



In order to decrease the heat load on the crystal, one can opt for the exploitation of a two-chicane self-seeding setup to increase the seed signal level against the SASE noise level, Fig. 10. The overall idea, first formulated in [65], originates from the observation that the signal impinging on the second crystal is characterized by a much narrower bandwidth than the competing SASE signal. It follows that the impinging signal level in the relevant frequency range around the seed is enhanced by a factor comparable with the ratio between SASE and seed bandwidths.

Simulations made for the European XFEL [66] showed that in the case of electron bunches [67] of 100 pC with the nominal peak current of 5 kA at 17.5 GeV the number of segments in the first and in the second part of the undulator can be reduced down to five and four, Fig. 10, compared to an optimum of seven for a single-chicane scheme (like that in Fig. 1) thus reducing the intensity of the X-ray pulses impinging on the crystals⁵ of about an order of magnitude, well below 3 μJ , see Fig. 11 and Fig. 12. In case of higher charges, corresponding to longer bunches, the energy impinging on the crystals increases too, but the advantage due to a better signal-to-noise ratio remains.

5 Self-seeding and tapering

The most promising way to increase the output power of the SASE radiation is by tapering the magnetic field of the undulator. Tapering [20, 21, 22, 23, 24, 25, 26] consists of a slow reduction of the field strength of the undulator in order to preserve the resonance wavelength as the kinetic energy of the electrons decreases due to the FEL process. The strong radiation field produces a ponderomotive well, which is deep enough to trap electrons. The radiation produced by these captured particles increases the depth of the ponderomotive well, so that electrons are effectively decelerated, leading to an increase of the radiated power. Tapering thus results in a much higher output power compared to the case of a uniform undulator. The undulator taper could be simply implemented as a step taper from one undulator segment to the next. The magnetic field tapering is provided by changing the undulator gap. A further increase in power is achievable by starting the FEL process from a monochromatic seed, rather than from noise [24, 25, 26]. The reason is the higher degree of longitudinal coherence of the radiation in the seeded case, which involves, with tapering, a large portion of the bunch in the energy-wavelength synchronism. In other words, in contrast to SASE, while slipping with respect to radiation, the electron beam experiences a well-behaved ponderomotive potential. It is therefore very natural to exploit self-seeding together with tapering [27, 28, 29, 30, 31]. In order to be most effective, tapering begins slightly before saturation, and requires the availability of a relatively long undulator. This situation will be realized in practice at the European XFEL.

Fig. 13 and Fig. 14 illustrate simulations of the output from the setup described in the previous paragraph, that is an HXRSS double-chicane setup at the SASE2 line of the European XFEL. In particular, the tapering law and the output energy, power and spectrum are shown. It seems feasible to produce TW-level pulses, owing to the length of the undulators at European XFEL.

Physical effects involved in the description of the tapering efficiency including 3D effects were studied in [32], where the Kroll-Morton-Rosenbluth 1D theory [22] is extended to a 3D model. In that paper an empirical law is also discussed, which

⁵ For high-repetition rate applications, at the European XFEL it was actually decided to install two-chicanes HXRSS setups.

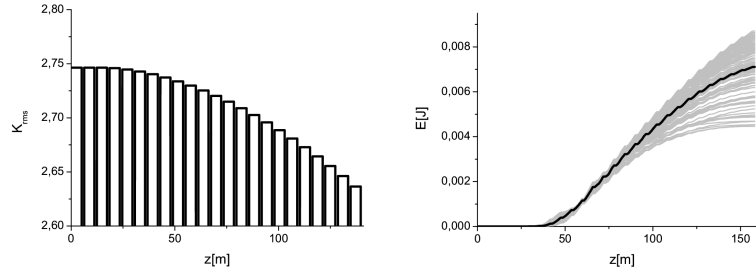
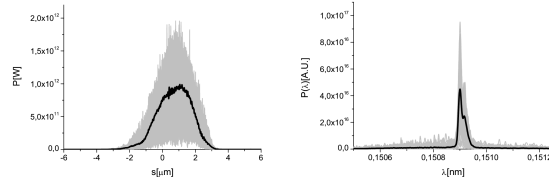


Fig. 13 Tapering law (left) and energy as a function of the position inside the undulator (right). These plots refer to the two-chicane HXRSS scheme for the SASE2 line at the European XFEL illustrated in Fig. 10. The grey lines represent 100 simulations, with different initial shot-noise conditions. The black line is the average over these 100 shots.

Fig. 14 Output power (left) and spectrum (right). These plots refer to the two-chicane HXRSS scheme for the SASE2 line at the European XFEL illustrated in Fig. 10. The grey lines represent 100 simulations, with different initial shot-noise conditions. The black line is the average over these 100 shots.



can be used to optimize the energy exchange between electrons and electromagnetic field, and which was also used to optimize the taper in the example discussed in this section. Indicating with z_0 the initial tapering position and with α and β two fitting parameters, one expresses the undulator parameter along the output undulator as $K(z) = K(z_0) [1 - \alpha(z - z_0)^\beta]$. Further studies of the dependency of the taper efficiency on the electron beam transverse profile can be found in [68]. In particular, in that reference it is shown that uniform and parabolic shapes lead to an increase in the total radiation power in comparison with a transverse Gaussian shape, and to a reduction of the energy in sidebands. Both at the LCLS [69, 12] and SACLA [16], experimental tapering studies are underway for HXRSS and SXRSS setups in order to optimize the output radiation power.

6 Summary and conclusions

Self-seeding [4] is an active filtering technique for X-ray FELs that allows for the production of nearly Fourier-limited photon pulses with a reduced spectral bandwidth, compared to conventional SASE. When coupled to tapering, self-seeding techniques promise an even larger increase in brightness. While the basic physics

concepts involved in the description of self-seeding are independent of the spectral range, the actual realization of a setup varies a lot from the hard X-ray range, roughly defined as starting from about 3 keV, where crystal reflections are available, to the soft X-ray range around 1 keV. In both cases considerable efforts have been undertaken to minimize the setup dimensions, resulting in the proposal of a wake monochromator setup in the hard X-ray case [14], and a grating monochromator setup in the soft X-ray range [56, 57].

HXRSS setups have been installed and successfully commissioned at the LCLS [15] and SACLA [16], while a SXRSS setup is now working at the LCLS [12]. Experimental demonstration showed that self-seeding techniques are much more sensitive to the characteristics of the electron beam and to its stability, especially the energy jitter, compared to usual SASE. This has triggered attention on the optimization of the electron beam quality both at the LCLS [48] and SACLA [16]). In the hard X-ray range, the presence of many usable reflections has enabled the possibility to deliver multiple colors within the FEL bandwidth, just by proper positioning of the crystal [49].

Finally, whenever high-repetition rate is required, as in the case of the European XFEL or the LCLS-II, special attention is needed to limit the heat load on the crystal. Excess heat can be removed by cooling techniques. However, regardless of the cooling technique considered, one can use a two-chicane scheme to increase the signal-to-noise figure of a factor roughly given by the ratio of the SASE bandwidth to the seeded bandwidth.

For the future, many facilities are considering the implementation of self-seeding setups [70, 18, 19, 63, 71] above the 1 keV energy range. This technique is particularly appealing, given its relative simplicity, especially in the harder X-ray region. Self-seeding has realized the dream of (nearly) single-mode FEL radiation pulses in the short X-ray wavelength region.

Acknowledgements

I am grateful Evgeni Saldin (DESY) for careful reading of the manuscript and his continuous advice. I thank Serguei Molodtsov (European XFEL) for his interest in this work and his support.

Appendix. Theory of Soft X-ray Self-Seeding (SXRSS) with gas monochromator

The self-seeding technique was originally devised for the soft X-ray range. Even before the successful demonstration of the HXRSS setup at the LCLS [15] and at SACLA [16], the community recognized the need for an extension of the method to the soft X-ray range. In fact, the use of diamond crystals to implement wake

monochromators is unfortunately limited by the distance between atomic layers to X-ray energies above 3 keV. However, as discussed in section 3.1, the general principle of operation of a wake monochromator does not include exclusively the use of crystals. In fact, the overall technique, albeit difficult to implement without crystals, is only based on the possibility of obtaining narrow bandwidth ‘holes’ in the SASE FEL spectrum. In [72] a possible extension of the wake monochromator method to the soft X-ray range was presented, based on a gas cell filled with resonantly absorbing gas. If the transmittance spectrum in the gas exhibits an absorbing resonance with a bandwidth narrow enough for seeding then, similarly to the hard X-ray case, the temporal waveform of the transmitted radiation pulse is characterized by a long monochromatic wake. In other words, the FEL pulse forces the gas atoms to oscillate in a way consistent with a forward-propagating, monochromatic radiation beam.

In [72] it was proposed to take advantage of autoionizing resonances in rare gases to seed in the XUV/soft X-ray range. The phenomenon of autoionization is well-known in literature (see for instance textbooks like [73], [74] and references therein). Here we consider Helium atomic gas as an example⁶. Autoionizing resonances result from the decay of doubly excited Rydberg states He^* into the continuum, i.e. $\text{He}^* \rightarrow \text{He}^+ + e^-$. Since the continuum can also be reached by direct photoionization, both paths add coherently, giving rise to interference. This interference is related to the typical Fano line shape for the cross-section as a function of energy [75]. Here we only report the cross section for these series, which can be modeled by the expression [76]:

$$\sigma(\lambda, q, \Gamma) = \sigma_b(\lambda) \left(\frac{(\sum_{n=2}^{\infty} (q_n / \mathcal{E}_n) + 1)^2}{\sum_{n=2}^{\infty} (1 / \mathcal{E}_n)^2 + 1} \right), \quad (8)$$

where the energy-dependent background cross-section expressed in Megabarn ($1\text{Mb} = 10^{-18} \text{cm}^2$) is given by $\sigma_b(\lambda) = -0.05504 - 1.3624 \cdot 10^{-4} \lambda + 3.3822 \cdot 10^{-5} \lambda^2$, with λ the radiation wavelength in Angstrom units, while the reduced energy \mathcal{E}_n is defined as $\mathcal{E}_n = 2(E_{R_n} - hc/\lambda)/\Gamma_n$. The asymmetry index q_n , the energy of the n th resonance E_{R_n} , the resonance width Γ_n , can be found, for example, in [76]. Using these parameters, the cross-sections for the series ($sp, 2n+$) can be calculated from Eq. (8) and are shown in Fig. 15 as a function of the energy. The photoabsorption cross section σ is linked to the light attenuation through a gas medium of column density $n_0 l$, where l is the length of the cell and n_0 the gas density.

If a monochromatic electromagnetic pulse of intensity I_0 and frequency ω impinges on a cell of length l , filled with a gas with density n_0 , the transmitted intensity obeys the Beer-Lambert law

$$I(\omega) = I_0 \exp[-n_0 l \sigma(\omega)]. \quad (9)$$

⁶ Other gas choices are possible. For example, in [72] also Neon was considered for photon energies between 45 eV and 49 eV, and Argon for a photon energy of 28.5 eV.

Fig. 15 Fano profiles for the $(sp, 2n+)$ autoionizing series of Helium. The cross-sections are calculated following [76].

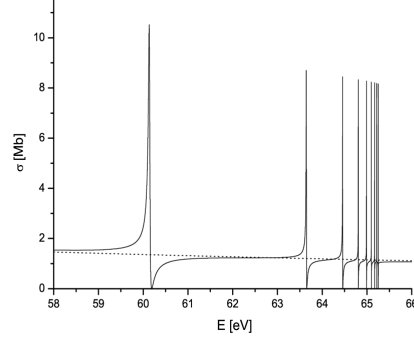
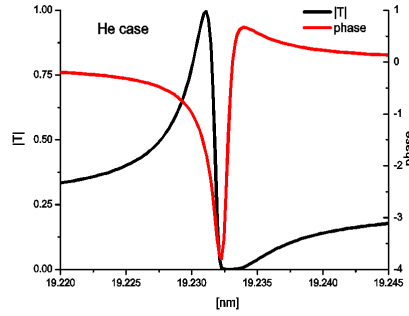


Fig. 16 Modulus and phase of the transmissivity of Helium around the $n = 4$ line of the $(sp, 2n+)$ 1P_0 Rydberg series. The phase is recovered from the modulus with the help of the Kramers-Kronig relation according to Eq. (2).



As a result, by comparison with Eq. (9), the modulus of the transmissivity can be defined as $|T| = \exp[-n_0 l \sigma(\omega)/2]$. We may then write

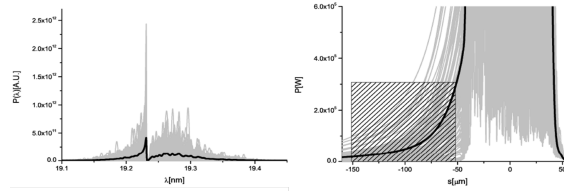
$$\ln[T(\omega)] = \ln[|T(\omega)|] + i\Phi(\omega) = -n_0 l \sigma(\omega)/2 + i\Phi(\omega). \quad (10)$$

In the following we will choose a column density equal to $n_0 l = 10^{18} \text{ cm}^{-2}$, and we will restrict ourselves to the third ($n = 4$) resonant line of the $(sp, 2n+)$ 1P_0 Rydberg series for Helium. The cross-sections can be calculated with the help of Eq. (8). This fixes $|T|$ as a function of frequency. Then, following the general treatment in section 2, one can use Kramers-Kronig relation to recover the phase according to Eq. (2). The final result in terms of modulus and phase of the transmissivity T is shown in Fig. 16.

The filter described here⁷ can substitute the single crystal considered in the previous section for HXRSS. Therefore, the self-seeding setup just consists of the gas cell, to be filled with noble gas, and a short magnetic chicane. Fig. 17 shows an example of an FEL pulse filtered with the help of the gas monochromator. It refers

⁷ Different choices of the column density can lead instead to a band-pass filter, as discussed in [77] and actually realized in [78].

Fig. 17 Example of filtered SASE spectrum (left) and power (right) after a gas monochromator. The highlighted part of the right plot illustrates the wake. Even if it is different from that created by a crystal monochromator, it can be used as a seed exactly in the same way.



to a case study described in detail in [72]. Note that the shape of the wake is now completely different compared to the case of a diamond wake monochromator, but this is not important. In fact, the radiation power within the wake is much larger than the equivalent shot noise power in the electron bunch. therefore, it can be used as seed pulse.

The scheme discussed here, while remaining of theoretical interest, has several practical limitations. In fact, although the availability of a series of resonances and of different gases allows for seeding at different frequencies, the scheme lacks continuous tunability. Moreover, not all resonances can be used. In fact, they should be located further away from each other than the FEL bandwidth, and be narrow enough to guarantee the applicability of the temporal windowing process. In the case in Fig. 17, a rule of thumb requires a width narrower than 10 meV. In this case, for example, only the second and the third Helium lines studied here are suitable, as the first one is too wide, and starting from the fourth there is not enough separation⁸.

Finally, it should be stressed that the autoionizing resonance that we considered here as an example is in the 20-nm range, which is outside the region of interest of the self-seeding technique. A possible way to generate shorter wavelengths, down to the few-nanometer range, is to use a two-stage output undulator, with the second stage resonant to one of the harmonics of the first one [72]. In this way, the second part of the output undulator acts effectively as a radiator, where one exploits the non-linear bunching present in the electron beam.

References

1. E. Saldin, E. Schneidmiller and M. Yurkov, ‘The Physics of Free-Electron Lasers’, Springer (1999).
2. A. Kondratenko and E. Saldin, Part. Accel. 10, 207 (1980).
3. R. Bonifacio, C. Pellegrini and L. Narducci, Opt. Commun. 50, 373 (1984).
4. J. Feldhaus et al., Optics. Comm. 140, 341 (1997).
5. P. Emma, et al., Nature Photon. 4, 641 (2010).
6. T. Ishikawa, et al., Nat. Photonics 6, 540 (2012).

⁸ Such smaller separation may be useful at VUV facilities, in order to seed at multiple photon wavelengths with separation of a fraction of eV, similarly to what has been proposed in [50].

7. M. Altarelli, et al. (Eds.) XFEL, The European X-ray Free-Electron Laser, Technical Design Report, DESY 2006-097, fulltext available at <http://arxiv.org/> (2006).
8. I. S. Ko and J.-H. Han, 'Current status of PAL-XFEL', in Proceedings of the 27th Linear Accelerator Conference, Geneva, Switzerland, MOIOB04 (2014).
9. SwissFEL Conceptual Design Report, PSI Bericht Nr. 10-04 (2012).
10. Ackermann et al. Nat. Photonics 1, 33642 (2007).
11. R. Treusch, W. Brefeld, J. Feldhaus and U Hahn, Ann. report 2001 'The seeding project for the FEL in TTF phase II' (2001).
12. D. Ratner et al., Phys. Rev. Letters, 114, 054801 (2015).
13. E. Saldin, E. Schneidmiller, Yu. Shvyd'ko and M. Yurkov, NIM A 475, 357 (2001).
14. Geloni, G., Kocharyan V., and Saldin, E., Journal of Modern Optics, 58, 1391 (2011).
15. J. Amann et al., Nature Photonics 6, 693 (2012).
16. T. Inagaki et al., 'Hard X-Ray Self-Seeding setup and results at SACLA', Proceedings of the 2014 FEL Conference, Basel, Switzerland, TUC01 (2014).
17. XFELSEED, 'Design and construction of Hard X-ray Self-Seeding Setups for the European XFEL', Project approved in the framework of the coordinated German-Russian call for proposals 'Ioffe-Roentgen Institute' (2014).
18. H.-S. Kang, 'Current status of PAL-XFEL project', Proceedings of the 2014 FEL Conference, Basel, Switzerland, WEODB103 (2014).
19. E. Prat and S. Reiche, 'Self-seeding design for SwissFEL', Proceedings of the 2013 FEL Conference, New York, USA, WEPSO5 (2013).
20. A. Lin and J.M. Dawson, Phys. Rev. Lett. 42, 2172 (1986).
21. P. Sprangle, C.M. Tang and W.M. Manheimer, Phys. Rev. Lett. 43 1932 (1979).
22. N.M. Kroll, P. Morton and M.N. Rosenbluth, IEEE J. Quantum Electron., QE-17, 1436 (1981).
23. T.J. Orzechowski et al., Phys. Rev. Lett. 57, 2172 (1986).
24. W. Fawley et al., NIM A 483, 537 (2002).
25. M. Cornacchia et al., J. Synchrotron Rad. 11, 227 (2004).
26. X. Wang et al., Phys. Rev. Letters 103, 154801 (2009).
27. G. Geloni, V. Kocharyan and E. Saldin, 'Scheme for generation of fully coherent, TW power level hard x-ray pulses from baseline undulators at the European XFEL', DESY 10-108, fulltext available at <http://arxiv.org/> (2010).
28. J. Wu et al., 'Staged self-seeding scheme for narrow bandwidth , ultra-short X-ray harmonic generation free electron laser at LCLS', Proceedings of the 2010 FEL Conference, Malmo, Sweden, TUPB08 (2010).
29. G Geloni, V. Kocharyan, and E. Saldin, 'Production of transform-limited X-ray pulses through self-seeding at the European X-ray FEL', DESY 11-165, fulltext available at <http://arxiv.org/> (2011).
30. W.M. Fawley et al., 'Toward TW-level LCLS radiation pulses', Proceedings of the FEL 2011 Conference, Shanghai, China, TUOA4 (2011).
31. J. Wu et al., 'Simulation of the Hard X-ray Self-seeding FEL at LCLS', Proceedings of the 2011 FEL Conference, Shanghai, China, MOPB09 (2011).
32. Y. Jiao et al. Phys. Rev. ST Accel. Beams 15, 050704 (2012).
33. G. Geloni, V. Kocharyan and E. Saldin, 'Scheme for generation of highly monochromatic X-rays from a baseline XFEL undulator', DESY 10-033, fulltext available at <http://arxiv.org/> (2010).
34. Y. Ding, Z. Huang and R. Ruth, Phys. Rev. ST Accel. Beams, 13, 060703 (2010).
35. O. Grimm, K. Klose and S. Schreiber, 'Double-pulse Generation with the FLASH Injector Laser for Pump-Probe Experiments', Proceedings of EPAC 2006 Conference, Edimburgh THPCH150 (2006).
36. A. Marinelli et al., Nature Communications DOI: 10.1038/ncomms7369 (2015).
37. V. Lucarini et al., 'Kramers-Kronig relations in optical materials research', Springer, (2004).
38. H. A. Kramers, 'La diffusion de la lumiere par les atomes', in Atti del Congresso Internazionale dei Fisici, 2, 545 (1927).

39. R. de L. Kronig, *J. Opt. Soc. Am.* 12, 547557 (1926).
40. J. Toll, *Phys. Rev.* 104, 6 (1956).
41. R. R. Lindberg and Y.V. Shvydko, *Phys. Rev. ST Accel. Beams* 15, 050706 (2012).
42. Y. Shvydko and R. Lindberg, *Phys. Rev. ST Accel. Beams* 15, 100702 (2012).
43. G. Geloni et al., 'Wake monochromator in asymmetric and symmetric Bragg and Laue geometry for self-seeding the European XFEL', DESY 13-013, fulltext available at <http://arxiv.org/> (2013).
44. D. Shu et al., *Journal of Physics: Conference Series* 425, 052004 (2013).
45. G. Geloni, V. Kocharyan and E. Saldin, 'Cost-effective way to enhance the capabilities of the LCLS baseline', DESY 10-133, fulltext available at <http://arxiv.org/> (2010).
46. Zhu, D. et al., *Appl. Phys. Lett.* 101, 034103 (2012).
47. F.-J. Decker et al., 'Improving and maintaining FEL beam stability of the LCLS', Proceedings of the 2014 IPAC Conference, Dresden, Germany, THPRO035 (2014).
48. J. Welch, 'Seeding at LCLS FEL', Joint DESY and University of Hamburg Accelerator Physics Seminar, Feb. 5, 2013, DESY Hamburg, <http://www-mpy.desy.de/AccPhySemDESY/y2013/JWelch.SeedingAtLCLS.pdf> (2013).
49. A. Lutman et al., *Phys. Rev. Lett.* 113, 254801 (2014).
50. G. Geloni, V. Kocharyan and E. Saldin, *Opt. Commun.* 284, 3348 (2011).
51. G. Geloni, V. Kocharyan and E. Saldin, 'Ultrafast X-ray pulse measurement method', DESY 10-008, fulltext available at <http://arxiv.org/> (2008).
52. Y. Ding et al., *Phys. Rev. Lett.* 109, 254802 (2012).
53. A. A. Lutman et al., *Phys. Rev. Letters* 110, 134801 (2013).
54. T. Hara et al., *Nature Communications*, DOI: 10.1038/ncomms3919 (2013).
55. G. Geloni, V. Kocharyan, and E. Saldin, 'Scheme for femtosecond-resolution pump-probe experiments at XFELs with two-color ten GW-level X-ray pulses', DESY-10-004, fulltext available at <http://arxiv.org/> (2010).
56. Y. Feng, J. Hastings, P. Heimann, M. Rowen, J. Krzywinski, and J. Wu, 'X-ray Optics for soft X-ray self-seeding the LCLS-II', Proceedings of the 2010 FEL conference, Malmo, Sweden, TUBP10 (2010).
57. Y. Feng et al., 'System design for self-seeding the LCLS at soft X-ray energies', Proceedings of the 2012 FEL Conference, Nara, Japan, TUOBI01 (2012).
58. S. Serkez et al., *Phys. Rev. ST AB* 18, 030708 (2015).
59. P. Emma et al. *Phys. Rev. Letters* 92, 074801-1 (2004).
60. P. Emma, M. Borland and Z. Huang, 'Femtosecond and Subfemtosecond X-Ray Pulses from a Self-Amplified Spontaneous-EmissionBased Free-Electron Laser', Proceedings of the 2004 FEL Conference, Trieste, Italy, TUBIS01, (2004).
61. A. Lutman et al., *J. Phys. A: Math. Theor.* 42, 045202 (2009).
62. G. Geloni, V. Kocharyan and E. Saldin, 'Self-Seeded Operation of the LCLS Hard X-ray FEL in the Long-Bunch Mode', DESY 10-239, fulltext available at <http://arxiv.org/> (2010).
63. T. Raubenheimer, 'The LCLS-II, a New FEL Facility at SLAC', Proceedings of the 2014 FEL Conference, Basel, Switzerland, WEB001 (2014).
64. H. Sinn, Private Communication (2012).
65. G. Geloni, V. Kocharyan and E. Saldin, 'Production of transform-limited X-ray pulses through self-seeding at the European X-ray FEL', DESY 11-165, fulltext available at <http://arxiv.org/> (2011).
66. G. Geloni, E. Saldin, V. Kocharyan, unpublished (2012).
67. I. Zagorodnov, 'Beam dynamics simulations for XFEL', <http://www.desy.de/fel-beam/s2e/> (2011).
68. C. Emma et al., *Phys. Rev. ST AB* 17, 110701 (2014).
69. J. Wu, 'Recent experimental and theoretical results on tapering for high intensity extraction', Workshop on Advanced X-ray FEL Development, Hamburg, Germany, http://www.xfel.eu/events/workshops/2014/workshop_on_advanced_x_ray_fel_development/ (2014).

70. 'Design and construction of Hard X-ray Self-Seeding Setups for the European XFEL (XFELSEED)', coordinated Russian-German project in the framework of the 'Ioffe-Röntgen' cooperation (2014).
71. F. Curbis et al., 'Simulation studies for an X-ray FEL based on an extension of the MAX IV linac', Proceedings of the 2013 FEL Conference, New York, USA, WEP007 (2013).
72. G. Geloni, V. Kocharyan, and E. Saldin, 'Self-seeding scheme with gas monochromator for narrow-bandwidth soft X-ray FELs', DESY 11-049, fulltext available at <http://arxiv.org/> (2011).
73. A. Bohm, 'Quantum Mechanics', Springer (1979).
74. U. Fano and A.R.P. Rau, 'Atomic collisions and spectra', Academic Press (1986).
75. U. Fano, Phys. Rev. 124, 6 (1961).
76. H. Morgan and D. Hederer, Phys. Rev. A 29, 4 (1984).
77. G. Geloni, V. Kocharyan, and E. Saldin, 'Gas-filled cell as a narrow bandwidth bandpass filter in the VUV wavelength range', DESY 11-055, fulltext available at <http://arxiv.org/> (2011).
78. J. Harries et al., J. Phys. B: At. Mol. Opt. Phys. 46, 164021 (2013).