applied sciences

MDPI

Article

Application of Matched-Filter Concepts to Unbiased Selection of Data in Pump-Probe Experiments with Free Electron Lasers

Carlo Callegari [1,2,*,†], Tsukasa Takanashi [3,†], Hironobu Fukuzawa [3], Koji Motomura [3,‡], Denys Iablonskyi [3], Yoshiaki Kumagai [3,§], Subhendu Mondal [3], Tetsuya Tachibana [3], Kiyonobu Nagaya [4], Toshiyuki Nishiyama [4], Kenji Matsunami [4], Per Johnsson [5], Paolo Piseri [6], Giuseppe Sansone [7,8], Antoine Dubrouil [7,‖], Maurizio Reduzzi [7], Paolo Carpeggiani [7], Caterina Vozzi [7], Michele Devetta [7], Davide Faccialà [7], Francesca Calegari [7,9], Mattea Carmen Castrovilli [7], Marcello Coreno [1,2], Michele Alagia [10], Bernd Schütte [11], Nora Berrah [12], Oksana Plekan [1], Paola Finetti [1], Eugenio Ferrari [1,¶], Kevin Charles Prince [1,10] and Kiyoshi Ueda [3]

1 Elettra-Sincrotrone Trieste S.C.p.A., Strada Statale 14–km 163.5 in Area Science Park, 34149 Basovizza, Trieste, Italy; marcello.coreno@elettra.eu (M.C.); oksana.plekan@elettra.eu (O.P.); paola.finetti12@gmail.com (P.F.); eugenio.ferrari@psi.ch (E.F.); kevin.prince@elettra.eu (K.C.P.)
2 CNR-ISM, Area Science Park, 34149 Basovizza, Trieste, Italy
3 Institute of Multidisciplinary Research for Advanced Materials, Tohoku University, Sendai 980-8577, Japan; tsukasat@mail.tagen.tohoku.ac.jp (Ts.T.); fukuzawa@tagen.tohoku.ac.jp (H.F.); motomura@spring8.or.jp (Ko.M.); denys@tagen.tohoku.ac.jp (D.I.); kumagay@anl.gov (Y.K.); justsm@gmail.com (S.M.); tach.ymn@gmail.com (Te.T.); ueda@tagen.tohoku.ac.jp (K.U.)
4 Department of Physics, Graduate School of Science, Kyoto University, Kyoto 606-8502, Japan; nagaya@scphys.kyoto-u.ac.jp (K.N.); t-nishiyama@scphys.kyoto-u.ac.jp (T.N.); matsunami.yaolab.kyoto@gmail.com (Ke.M.)
5 Department of Physics, Lund University, P.O. Box 118, 22100 Lund, Sweden; per.johnsson@fysik.lth.se
6 CIMAINA and Dipartimento di Fisica, Università degli Studi di Milano, Via Celoria 16, 20133 Milano, Italy; paolo.piseri@unimi.it
7 CNR-IFN, Piazza Leonardo da Vinci 32, 20133 Milano, Italy; giuseppe.sansone@physik.uni-freiburg.de (G.S.); dubrouil@femtoeasy.eu (A.D.); mauriziobattista.reduzzi@polimi.it (M.R.); paolo.carpeggiani@tuwien.ac.at (P.C.); caterina.vozzi@ifn.cnr.it (C.V.); michele.devetta@mail.polimi.it (M.D.); davide.facciala@polimi.it (D.F.); francesca.calegari@desy.de (F.C.); matteac@gmail.com (M.C.C.)
8 Physikalisches Institut, Albert-Ludwigs-Universität, 79104 Freiburg, Germany
9 Center for Free-Electron Laser Science, DESY, Notkestrasse 85, 22607 Hamburg, Germany
10 CNR-IOM, Area Science Park, 34149 Basovizza, Trieste, Italy; alagiam@elettra.eu
11 Max-Born-Institut, Max-Born-Strasse 2 A, 12489 Berlin, Germany; schuette@mbi-berlin.de
12 Department of Physics, University of Connecticut, 2152 Hillside Road, Storrs, CT 06269, USA; nora.berrah@uconn.edu
* Correspondence: carlo.callegari@elettra.eu; Tel.: +39-040-375-8844
† These authors contributed equally to this work.
‡ Current address: RIKEN SPring-8 Center, Kouto, Sayo, Hyogo 679-5148, Japan.
§ Current address: Argonne National Laboratory, 9700 S. Cass Avenue, Argonne, IL 60439, USA.
‖ Current address: Femto Easy, Parc scientifique Laseris 1, 33114 Le Barp, France.
¶ Current address: Particle Accelerator Physics Laboratory, École Polytechnique Fédérale de Lausanne EPFL, CH-1015 Lausanne, Switzerland.

Academic Editor: Malte Kaluza
Received: 27 April 2017; Accepted: 9 June 2017; Published: 16 June 2017

Abstract: Pump-probe experiments are commonly used at Free Electron Lasers (FEL) to elucidate the femtosecond dynamics of atoms, molecules, clusters, liquids and solids. Maximizing the signal-to-noise ratio of the measurements is often a primary need of the experiment, and the aggregation of repeated, rapid, scans of the pump-probe delay is preferable to a single long-lasting scan.

The limited availability of beamtime makes it impractical to repeat measurements indiscriminately, and the large, rapid flow of single-shot data that need to be processed and aggregated into a dataset, makes it difficult to assess the quality of a measurement in real time. In post-analysis it is then necessary to devise unbiased criteria to select or reject datasets, and to assign the weight with which they enter the analysis. One such case was the measurement of the lifetime of Intermolecular Coulombic Decay in the weakly-bound neon dimer. We report on the method we used to accomplish this goal for the pump-probe delay scans that constitute the core of the measurement; namely we report on the use of simple auto- and cross-correlation techniques based on the general concept of "matched filter". We are able to unambiguously assess the signal-to-noise ratio (*SNR*) of each scan, which then becomes the weight with which a scan enters the average of multiple scans. We also observe a clear gap in the values of *SNR*, and we discard all the scans below a *SNR* of 0.45. We are able to generate an average delay scan profile, suitable for further analysis: in our previous work we used it for comparison with theory. Here we argue that the method is sufficiently simple and devoid of human action to be applicable not only in post-analysis, but also for the real-time assessment of the quality of a dataset.

Keywords: correlation; matched filter; pump-probe; free electron laser; data processing; statistical weight

1. Introduction

Recently we described the results of a pump-probe experiment, in which the lifetimes of doubly excited states of neon dimers were measured [1]. The dimers were excited by absorption of two EUV photons from the Free Electron Laser (FEL) FERMI-1 [2] and probed via ionization by a UV laser pulse.

The excited dimers decayed by Interatomic Coulombic Decay to stable dimer cations Ne_2^+, which were detected by a time-of-flight (TOF) mass spectrometer. Ionization by the UV pulse led to a repulsive state of the dimer, which dissociated, so that the yield of dimer ions was reduced, and the yield of Ne^+ increased. The dimer sample used in that work was very dilute with a large atomic background, as it was produced in a supersonic expansion of neon gas, with a yield of about 1%. The data were therefore noisy, and a substantial effort was needed to analyze and filter them. In this paper, we describe the methods used for that analysis.

The data consisted of a number of temporal scans, each lasting approximately 30 min, in which the ion-TOF signal was measured as a function of the delay between pump and probe pulses, and these scans had poor signal-to-noise ratios. Simply averaging them did not give good results, probably because some of them suffered from poor FEL conditions or poor alignment of the pump and probe pulses. One should reject those scans that do not contribute significant signal, while avoiding human bias. A possible approach is to exploit the difference between signal and noise with respect to auto- and cross-correlation. This is an instance of the more general concept of "matched filter", i.e., the linear filter maximizing the signal-to-noise ratio (SNR) of a measured noisy sample [3]. Under the assumption of white noise, the matched filter is the complex-conjugate time-reversal of the signal sought, i.e., application of the matched filter returns the autocorrelation of the signal [4]. The concept is widely applied in signal processing, where one is primarily interested in extracting a burst of periodic signal from a noisy sample. As far as peak detection is concerned, a popular field of application is chromatography, but the general results obtained there apply equally well to our case; Ref. [5] explicitly discusses the use of a matched filter to determine amplitude and time shift of the peak being sought. We consider the use of auto- and cross-correlation for three purposes:

1. validate the scans to be included or excluded from averaging
2. determine the weight with which each scan should enter the average
3. determine, if desired, by how much to temporally shift a scan prior to averaging

2. Results

For a set of delays $\{d_i\}$, $i = 0..n$, let us consider two delay scans $\mathbf{R} = \{R_0, \ldots, R_n\}$ and $\mathbf{S} = \{S_0, \ldots, S_n\}$ each consisting of a sequence of Ne_2^+ ion-TOF signals, reduced as explained in Section 4, and padded with zeros outside of the delay range scanned. We will consider $\mathbf{R}$ our reference scan, which for convenience is assumed to be noise-free.

The definition of cross-correlation without normalization, $\Gamma_{\mathbf{RS},q}$ is:

$$\Gamma_{\mathbf{RS},q} = \sum_i R_i S_{i+q} \equiv (\mathbf{R} \star \mathbf{S})_q \tag{1}$$

Note that Γ is linear in the two sequences (scaling either one by a factor α scales Γ by the same factor), and that $\Gamma_{\mathbf{SS},q}$ is the autocorrelation of $\mathbf{S}$, which has a maximum for $q = 0$:

$$(\mathbf{S} \star \mathbf{S})_0 = \sum_i (S_i)^2 \tag{2}$$

We now assume that any scan $\mathbf{N} = \{N_0, \ldots, N_n\}$ consisting of pure noise has zero (negligible) cross-correlation with any other scan (including itself except, obviously, at zero-shift), thus:

$$(\mathbf{N} \star \mathbf{N})_q = \delta_{q0} \sum_i (N_i)^2 \tag{3}$$

$$(\mathbf{R} \star \mathbf{N})_q = 0 \tag{4}$$

with δ_{q0} the Kronecker delta; we use Equation (3) to define a variance $\sigma^2 = (\mathbf{N} \star \mathbf{N})_0 = \sum_i (N_i)^2$. Equations (3) and (4) strictly hold for white noise and infinite n; in a real-life situation we can expect that noise correlation just decays much faster than signal correlation: thus all scans whose autocorrelation is sharply peaked are probably pure noise, and will also have poor cross-correlation with the reference scan.

Any noisy scan $\mathbf{S}$ can be written in terms of the reference scan $\mathbf{R}$ and a pure-noise sequence $\mathbf{N}$ as: $\mathbf{S} = \alpha\mathbf{R} + \mathbf{N}$, with α a real number. In a worse-case scenario, S_i may be shifted (for simplicity by an integer index r), that is: $S_i = \alpha R_{i+r} + N_i$. Then:

$$(\mathbf{R} \star \mathbf{S})_q = \alpha(\mathbf{R} \star \mathbf{R})_{q-r} + (\mathbf{R} \star \mathbf{N})_q = \alpha(\mathbf{R} \star \mathbf{R})_{q-r} \tag{5}$$

note that $\mathbf{R} \star \mathbf{S}$ is proportional to $\mathbf{R} \star \mathbf{R}$ but shifted by r as expected; this is the argument invoked in Ref. [5] to associate the position of the maximum of the cross-correlation function to the shift of the peak being sought. In our work [1] it was not necessary to include a shift; while Equation (5) could be used to estimate the signal amplitude α, it is preferable to use the autocorrelation instead:

$$(\mathbf{S} \star \mathbf{S})_q = \alpha^2(\mathbf{R} \star \mathbf{R})_q + (\mathbf{N} \star \mathbf{N})_q = \alpha^2(\mathbf{R} \star \mathbf{R})_q + \sigma^2\delta_{q0} \tag{6}$$

(note that the shift by r cancels, here). Equation (6) tells us in particular that all autocorrelation sequences should be the same except for a scale factor α^2 and a sharp noise peak at $q = 0$. Visual inspection shows that this is true for a number of scans which we consider good; other scans exhibit a narrower, or structured, sometimes negative, autocorrelation, probably indicating correlated noise, i.e., a drift during the measurement; the rest only exhibits the noise peak, indicating no signal at all (Figure 1). From now on we will consider $(\mathbf{R} \star \mathbf{R})$ to be noiseless and, as empirically found, satisfactorily approximated by a gaussian; we derive the width of the gaussian from our best scan (Figure 1).

Our choice of a gaussian is purely empirical, and our method does not critically depend on it: to the extent that all good scans have the same shape of the autocorrelation function, Equation (6), the best fit parameters will be the same for all good scans, except for a scale factor for the height. What is important is that the fitting function provides a reasonable approximation of α in Equation (6), i.e., of the peak value of the signal component of the autocorrelation.

It is nevertheless instructive to discuss some limiting cases for the shape of the autocorrelation function: we begin by noting that the expected shape of a pump-probe signal such as that of our experiment is an exponential decay (we ignore the possible complication brought on by the presence of more than one decay constant) convoluted with the instrumental resolution (in our case a gaussian, coming from the finite duration of the FEL and UV pulses). The respective autocorrelation functions are an exponential and a gaussian. For the case of non-negligible instrumental resolution the autocorrelation will resemble a gaussian near $q = 0$, but will have broader wings decaying as $\exp(-|q|)$ rather than $\exp(-q^2)$. An early truncation of the scan will clip the wings of the autocorrelation function. In our case an early truncation does somewhat contribute to determining the shape of the autocorrelation function, but does so equally for all scans (the pump-probe delay was scanned in reverse, and point index $i = 0$ in Figure 1 consistently corresponds to the maximum value of the delay).

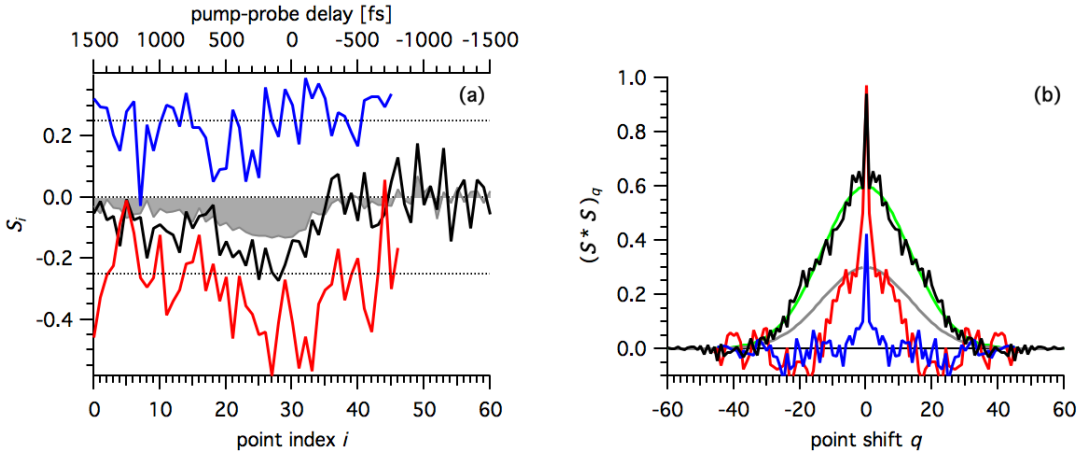

Figure 1. (**a**) Three representative delay scans and (**b**) their autocorrelation curves. We use the autocorrelation curves ($q = 0$ excluded) as a criterion to classify scans as good (black), drifting (red), or pure noise (blue). The autocorrelation curve of a good scan is well approximated by a gaussian (green). In panel (**a**), traces have been offset for clarity, and the shaded curve is the weighed average Equation (7); in panel (**b**), the green curve is a gaussian fit of the black autocorrelation trace; the gray curve is the same, scaled to the height of the red autocorrelation trace (white-noise peak excluded). The calculated signal-to-noise ratio (SNR) for the three scans are 1.30, 0.52, and 0.20.

Let us finally consider two possible sources of artefacts, namely a constant, or a linearly drifting, baseline (remembering that the scans are padded with zeros outside of the delay range scanned). This would primarily contribute a slowly decaying component to the autocorrelation signal (of width comparable to that of the scan itself). Let us however note that because the scans are acquired by rapid double-background subtraction (Equation (9) and Figure 4), we expect a complete baseline cancellation. Finally, a drifting signal would alter the shape of the peak and consequently of its autocorrelation function: assuming an even drift that causes a loss of the optimal experimental conditions (temporal or spatial overlap; quality of the focus; resonance wavelength) one can speculate that its main effect would be a reduction of the measured width.

3. Discussion

Given a set of scans $\left\{\mathbf{S}^{(k)}\right\}$, their weighted average is

$$\langle \mathbf{S} \rangle = \frac{\sum_k w_k \mathbf{S}^{(k)}}{\sum_k w_k} \tag{7}$$

$\langle \mathbf{S} \rangle$ is the quantity reported in our work ([1] note for the sake of exactness that in Figure 2a therein, the unity baseline was not subtracted); we want to determine which scans to include in the average, and their weights w_k.

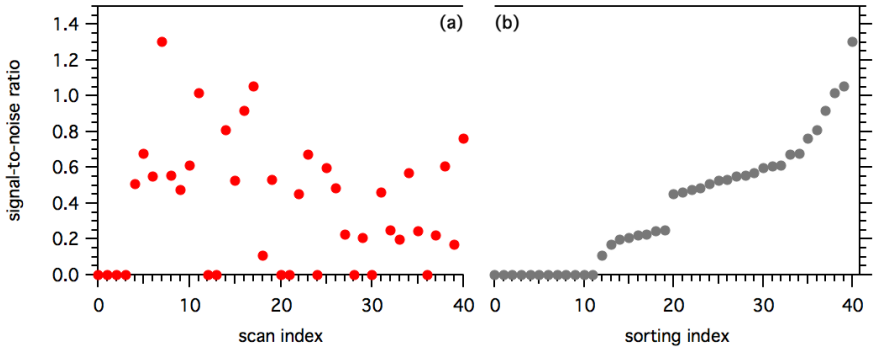

Figure 2. (**a**) Signal-to-noise ratio plotted versus scan index. (**b**) Same data as (**a**) sorted by increasing value, to highlight the gap between $SNR > 0.25$ and $SNR < 0.45$.

From Equation (6) and Figure 1b we can estimate the signal-to-noise ratio as:

$$SNR_k = \sqrt{\frac{(\mathbf{S}^{(k)} \star \mathbf{S}^{(k)})_0}{(\mathbf{N}^{(k)} \star \mathbf{N}^{(k)})_0} - 1} = \sqrt{\frac{\alpha_k^2 (\mathbf{R} \star \mathbf{R})_0}{(\mathbf{N}^{(k)} \star \mathbf{N}^{(k)})_0}} = \frac{\alpha_k}{\sigma_k} \tag{8}$$

A simple analysis of its trend over the course of the experiment reveals some regularities that we exploit to qualitatively classify scans, and to define a quantitative criterion that we adopt to accept or reject them. When the scans are ordered in the sequence they were acquired (Figure 2a) no obvious trend is visible for the SNR, but one does note a large number of scans with $SNR = 0$: they are either those for which a gaussian fit of the autocorrelation was not successful, or those which have been excluded *a priori* (e.g., because the scan was aborted). The ordering of the scans by increasing SNR (Figure 2b) reveals a gap between scans with $SNR < 0.25$ and $SNR > 0.45$. We cannot find an obvious reason for this gap; because this was one of the first resonant two-photon experiments performed at FERMI we can speculate a threshold behavior of some of the less-controlled parameters of the FEL (peak intensity; second harmonic content). In any case, we decided to use the condition $SNR > 0.45$ as a discriminant to include a scan in the average.

Let us now come to the weight with which each accepted scan enters the average. We show in Appendix A that the weight which maximizes the SNR of $\langle \mathbf{S} \rangle$ is $w_k = \alpha_k / \sigma_k^2 = SNR_k / \sigma_k$; in Ref. [1] we used $w_k = SNR_k$, which gives a slightly worse result, although the difference is not significant. We presume that the latter fact depends on the limited number of samples, the predominance of few of them (see Figure 2a), and the low dispersion of the σ_k. In both cases we observe an improvement of the signal-to-noise ratio of $\langle \mathbf{S} \rangle$ by a factor ≈2.25 relative to the best single scan.

Unfortunately at the time of the experiment we were not anticipating the need for this test, so we are not able to perform further checks on which parameters and events mostly affected our experiment. Likewise our primary goal was to identify a simple unbiased method of data selection, and we do not attempt a systematic analysis of its merits and limitations, for which we refer the reader to the vast existing literature ([6,7] and references therein); we simply note that the method is most easily applicable to the case of white noise and single- or well-separated peaks, although generalizations to other noise [6] or multiple peaks [8] have been discussed. Let us finally note that because the autocorrelation is the Fourier Transform of the power spectrum (Wiener-Khinchin Theorem), the same determination of the signal-to-noise ratio could have been accomplished by Fourier Transform of each scan (in the frequency domain, the white-noise component appears as a constant).

Despite its simplicity and limitations, we believe that our simple test can be further characterized and profitably applied in future experiments. Let us note that the original and most immediate application of the method addresses a situation common to the beginning of many experiments, namely the need of making a weak signal visible when the expected signal shape is not precisely known in advance.

4. Materials and Methods

4.1. Experimental

The experiment was performed at the Low Density Matter (LDM) beamline [9] at FERMI FEL-1 [2]. The EUV pulses had an average energy of 16 µJ, duration estimated between 60 and 80 fs FWHM, circular polarization, focal size 30 µm FWHM, and were tuned to the wavelength resonant with the two-photon doubly excited target state, 75.65 nm. The UV pulses had an energy of 35 µJ, estimated duration 200 fs, focal size 80 µm FWHM, fixed wavelength (261 nm). The Ne dimers were produced by adiabatic expansion of Ne gas at a pressure of 0.8 MPa and temperature of 190 K through a 100 µm nozzle. The target ions were detected by the time-of-flight mass spectrometer of the LDM endstation [10].

4.2. Data Acquisition

The Free Electron Laser FERMI currently operates at a repetition rate of 50 Hz (10 Hz at the time of the experiment). This is determined by the operation rate of the LINAC that accelerates the electron bunches used to generate the FEL pulses. Single-shot ion TOF traces (Figure 3) are acquired, tagged with a progressive integer ("bunchnumber") and stored along with a wide selection of experimental parameters and machine parameters, for post-processing. The relevant single-shot quantities for us are the FEL energy per pulse ($I0$), the UV energy per pulse (I_{UV}), and the integrated area a of the TOF peak of interest (specifically the $^{20}Ne_2^+$ peak).

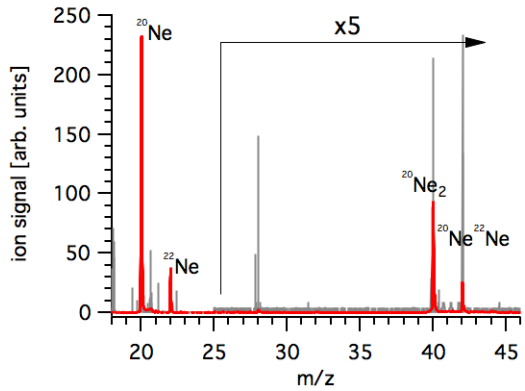

Figure 3. Single-shot ion-TOF trace (gray) and average over 12,200 shots (red). One in three shots is a blank shot (no gas sample), which is subtracted from the average to eliminate spurious signals, such as that at $m/z = 28$ in the single shot spectrum, due to residual nitrogen gas. Note that for $m/z > 25$ both spectra are magnified by a factor 5.

A preset number of single-shot values a_j at nominally identical conditions are acquired and aggregated into a datapoint A_i; one in three shots is a blank shot (no gas sample), which is subtracted from the average to eliminate spurious signals (Equation (9)). Although in reality the bunchnumber is never reset, i.e., each bunchnumber is unique since the inception of FERMI, for the purpose of this work we will think of it as spanning a range $j = \{1..m\}$ for each datapoint. A set of datapoints at a sequence of delay values constitutes a delay scan. Delay scans are averaged into a global delay scan, where each scan enters with the weight determined in Equation (8).

4.3. Data Reduction

From the single-shot values we build a datapoint as follows:

$$A_i = \frac{\sum\limits_{j=\text{gas}} a_j}{\sum\limits_{j=\text{gas}} (I0_j)^p} - \frac{\sum\limits_{j=\text{blank}} a_j}{\sum\limits_{j=\text{blank}} (I0_j)^p} \tag{9}$$

with the new index i labelling sets of nominally identical measurements at a series of pump-probe delays d_i. $\tilde{A}_i$ is the analogue of A_i with UV off. We consider equally spaced delays, so $d_i = \Delta t \times i \quad (i = 0..n)$. Let us note that single-shot data can be excluded from the sum by applying in post-analysis suitable filter conditions (e.g., a threshold on $I0$, or on the quality of the FEL spectrum); for this reason the number of single-shots aggregated into a datapoint is not strictly constant. For each i we define the quantity $S_i = A_i/\tilde{A}_i - 1$: this is the base datapoint of a delay scan (Figure 4). The integer p in Equation (9) is the normalization order; for the data presented in Ref. [1] we chose $p = 0$, i.e., no normalization. Let us also note that the intensity of the UV pulse was kept constant and is very stable, so it does not appear in the data processing. Let us finally note that at negative delays, and large positive delays we expect the UV to have no effect, i.e., $S_i = 0$.

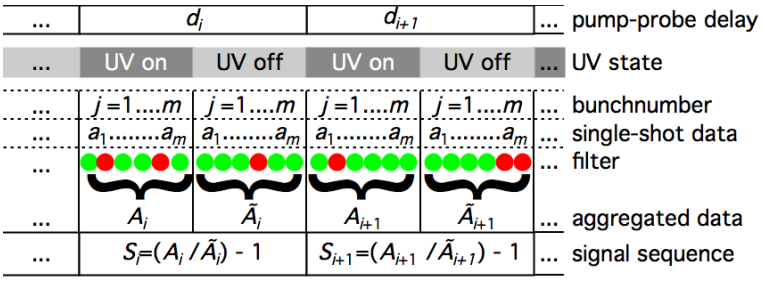

Figure 4. Schematic representation of the generation of a delay scan $\{S_i\}$, as a function of pump-probe delay $\{d_i\}$ from single-shot data a_j. The green (red) dots symbolize shots that pass (fail) the filter conditions.

5. Conclusions

Although rather simple, the method proposed has many advantages and applications. It provides a quick and reliable method to evaluate the quality of a measurement in real time, even when the shape of the expected signal is unknown. In fact, in the approximation of white noise, and of an unstructured pump-probe signal, it provides a first coarse estimate of the signal shape, and thus of the delay range that needs to be scanned. In post-analysis it provides an unbiased criterion to include each scan in an averaged measurement.

Acknowledgments: This work was supported by the X-ray Free Electron Laser Priority Strategy Program of the Ministry of Education, Culture, Sports, Science, and Technology of Japan (MEXT); by JSPS and CNR under the Japan - Italy Research Cooperative Program; by the Grant-in-Aid for the Global COE Program 'the Next Generation of Physics, Spun from Universality and Emergence' from the MEXT; by the Grants-in-Aid (No. 20310055 and No. 21244062) from the Japan Society for the Promotion of Science (JSPS); by the JSPS KAKENHI Grant Number JP 16J02270; by the Swedish Research Council and the Swedish Foundation for Strategic Research; by the ERC Starting Research Grant UDYNI No. 307964; by the European Union Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 641789 "MEDEA" (Molecular Electron Dynamics investigated by IntensE Fields and Attosecond Pulses); by the DOE-SC-BES under Award No. DE-SC0012376; and by the Italian Ministry of Education, Universities and Research (MIUR) (PRIN 2012 - NOXSS)

Author Contributions: K.U. conceived and designed the experiment; A.D., B.S., C.C., C.V., D.F., D.I., E.F., F.C., G.S., H.F., K.C.P., Ke.M., Ko.M., K.N., K.U., M.A., M.C., M.C.C, M.D., M.R., N.B., O.P., P.C., P.F., P.J., P.P., S.M., T.N., Ts.T., Te.T., Y.K. performed the experiments; C.C., H.F., Ko.M., Ts.T. analyzed the data; A.D., C.C., P.J., P.P. contributed analysis tools; C.C. and K.C.P. drafted the paper, which was completed in consultation with all authors.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

EUV	Extreme Ultra Violet
FEL	Free Electron Laser
FERMI	Free Electron laser Radiation for Multidisciplinary Investigations (this acronym identifies the Free Electron Laser facility in Trieste, Italy)
FWHM	Full Width at Half Maximum
I0	FEL energy per pulse
LINAC	Linear Particle Accelerator
SNR	Signal-to-noise Ratio
TOF	Time-of-flight
UV	Ultra Violet

Appendix A

We demonstrate that within a proportionality factor the optimum value of the weights appearing in Equation (7) is $w_j = \frac{\alpha_j}{\sigma_j^2}$; all three quantities α_j, σ_j, w_j are assumed positive; note that we use j to indicate a specific index, and tacitly replace it with k whenever it becomes a dummy index in a sum. We wish to maximize

$$\frac{\left(\sum_k w_k \alpha_k\right)^2}{\sum_k w_k^2 \sigma_k^2} - \lambda \left(1 - \sum_k w_k\right) \tag{A1}$$

where λ is a Lagrange multiplier. The condition of zero partial derivatives with respect to each of the w_j and of λ yields:

$$2a_j \left(\sum_k w_k \alpha_k\right) \left(\sum_k w_k^2 \sigma_k^2\right) - 2w_j \sigma_j^2 \left(\sum_k w_k \alpha_k\right)^2 + \lambda \left(\sum_k w_k^2 \sigma_k^2\right)^2 = 0 \tag{A2}$$

$$\sum_k w_k = 1 \tag{A3}$$

Multiplying (A2) by w_j and summing over j, one has:

$$2 \left(\sum_k w_k \alpha_k\right)^2 \left(\sum_k w_k^2 \sigma_k^2\right) - 2 \left(\sum_k w_k^2 \sigma_k^2\right) \left(\sum_k w_k \alpha_k\right)^2 + \lambda \left(\sum_k w_k\right) \left(\sum_k w_k^2 \sigma_k^2\right)^2 = 0 \tag{A4}$$

hence $\lambda = 0$; and Equation (A2), divided by $2\sigma_j^2 \neq 0$, simplifies to

$$\frac{a_j}{\sigma_j^2} \left(\sum_k w_k^2 \sigma_k^2\right) - w_j \left(\sum_k w_k \alpha_k\right) = 0 \tag{A5}$$

Summing over j, and using Equation (A3) yields

$$\sum_k w_k \alpha_k = \left(\sum_k \frac{a_k}{\sigma_k^2}\right) \left(\sum_k w_k^2 \sigma_k^2\right) \tag{A6}$$

which replaced in Equation (A5) finally gives

$$w_j = \frac{a_j / \sigma_j^2}{\sum_k a_k / \sigma_k^2} \tag{A7}$$

Note that in the special case when α_k is the same for all k, one has the familiar result that the minimum variance of an average is the one with $w_j = \sigma_j^{-2} / \sum_k \sigma_k^{-2}$ [11].

References

1. Takanashi, T.; Golubev, N.V.; Callegari, C.; Fukuzawa, H.; Motomura, K.; Iablonskyi, D.; Kumagai, Y.; Mondal, S.; Tachibana, T.; Nagaya, K.; et al. Time-Resolved Measurement of Interatomic Coulombic Decay Induced by Two-Photon Double Excitation of Ne_2. *Phys. Rev. Lett.* **2017**, *118*, 33202.
2. Allaria, E.; Appio, R.; Badano, L.; Barletta, W.A.; Bassanese, S.; Biedron, S.G.; Borga, A.; Busetto, E.; Castronovo, D.; Cinquegrana, P.; et al. Highly coherent and stable pulses from the FERMI seeded free-electron laser in the extreme ultraviolet. *Nat. Photonics* **2012**, *6*, 699–704.
3. Green, P. Preface to the matched filter issue. *IRE Trans. Inf. Theory* **1960**, *6*, 310.
4. North, D. An Analysis of the factors which determine signal/noise discrimination in pulsed-carrier systems. *Proc. IEEE* **1963**, *51*, 1016–1027.
5. Van den Heuvel, E.; van Malssen, K.; Smit, H. Optimal estimation of intensity of noisy peaks by matched filtering with application to chromatography. *Anal. Chim. Acta* **1990**, *235*, 343–353.
6. Smit, H. Specification and estimation of noisy analytical signals: Part I. Characterization, time invariant filtering and signal approximation. *Chemom. Intell. Lab. Syst.* **1990**, *8*, 15–27.
7. Smit, H. Specification and estimation of noisy analytical signals: Part II. Curve fitting, optimum filtering and uncertainty determination. *Chemom. Intell. Lab. Syst.* **1990**, *8*, 29–41.
8. Van den Bogaert, B.; Boelens, H.F.M.; Smit, H.C. Quantification of overlapping chromatographic peaks using a matched filter. *Chemom. Intell. Lab. Syst.* **1994**, *25*, 297–311.
9. Svetina, C.; Grazioli, C.; Mahne, N.; Raimondi, L.; Fava, C.; Zangrando, M.; Gerusina, S.; Alagia, M.; Avaldi, L.; Cautero, G.; et al. The Low Density Matter (LDM) beamline at FERMI: optical layout and first commissioning. *J. Synchrotron Radiat.* **2015**, *22*, 538–543.
10. Lyamayev, V.; Ovcharenko, Y.; Katzy, R.; Devetta, M.; Bruder, L.; LaForge, A.; Mudrich, M.; Person, U.; Stienkemeier, F.; Krikunova, M.; et al. A modular end-station for atomic, molecular, and cluster science at the low density matter beamline of FERMI@Elettra. *J. Phys. B* **2013**, *46*, 164007.
11. Bevington, P.R.; Robinson, D.K. *Data Reduction and Error Analysis for the Physical Sciences*, 2nd ed.; McGraw-Hill: New York, NY, USA, 1992.