Quantifying noise sources in the KSTAR 2014 Thomson Scattering system from the measured variation on electron temperature

T.-S. Oh, K.H. Kim, J.H. Lee, S.H. Lee, R. Scannell, A.R. Field, K. Cho, M.S. Bawa’aneh, and Y.-c. Ghim

Department of Nuclear and Quantum Engineering, KAIST, Daejeon, Korea
National Fusion Research Institute, Daejeon, Korea
Culham Centre for Fusion Energy, Culham Science Centre Abingdon, Oxfordshire OX14 3DB, U.K.
Department of Physics, Sogang University, Seoul, Korea
Department of Physics, The Hashemite University, Zarqa, Jordan
Department of Applied Mathematics and Sciences, Khalifa University, U.A.E.
Department of Physics, Yarmouk University, Irbid, Jordan

E-mail: ycghim@kaist.ac.kr

ABSTRACT: With the Thomson scattering (TS) system in KSTAR, temporal evolution of electron temperature ($T_e$) is estimated using a weighted look-up table method with fast sampling (1.25 or 2.5 GS/s) digitizers during the 2014 KSTAR campaign. Background noise level is used as a weighting parameter without considering the photon noise due to the absence of information on absolute photon counts detected by the TS system. Estimated electron temperature during a relatively quiescent discharge are scattered, i.e., 15% variation on $T_e$ with respect to its mean value. We find that this 15% variation on $T_e$ cannot be explained solely by the background noise level which leads us to include photon noise effects in our analysis. Using synthetic data, we have estimated the required photon noise level consistent with the observation and determined the dominant noise source in KSTAR TS system.

KEYWORDS: Plasma diagnostics - interferometry, spectroscopy and imaging; Analysis and statistical methods; Data processing methods

© 2016 IOP Publishing Ltd and Sissa Medialab srl
doi:10.1088/1748-0221/11/03/C03036
1 Introduction

Thomson scattering (TS) systems are widely used to measure the temperature ($T_e$) and density ($n_e$) of electrons in fusion devices. KSTAR Thomson scattering system consists of separate core and edge collection optics collecting the Thomson scattered photons which are transmitted to polychromators via optical fibres [1, 2]. After the band-pass filters, where figure. 1 shows an example of the measured filter functions, optical signals are converted to electrical signals and recorded with the newly installed fast sampling digitizers (NI PXIe-5160) with the sampling rate of either 1.25 or 2.5 GS/s, in addition to the existing charge-integrating (gating) digitizer.

We use a weighted look-up table method to estimate the time evolution of electron temperature while exclusively considering the uncertainty due to the background noise. Another possibly large source of uncertainty may originate from the photon (Poisson) noise. However, the photon noise could not be reflected in our analysis due to the lack of absolute photon counts. Note that the absolute photon counts can be estimated based on Rayleigh or rotational Raman calibration [3, 4], but such data have not been obtained with the fast sampling digitizers during the 2014 KSTAR campaign. Rayleigh calibration data are obtained only with the existing charge-integrating (gating) digitizers.

Estimated electron temperature shows scatters, i.e., 15% variation on $T_e$ with respect to its mean value. By using the synthetic data, we have found that this scattering on $T_e$ cannot be explained solely by the uncertainty due to the background noise. Our main goal of this paper is to estimate the ‘total’ noise level of the Thomson signal consistent with the observed variation on $T_e$ using the synthetic data. With the levels of estimated total noise and the measured background noise, we deduce the photon noise level assuming that photon noise and background noise are uncorrelated. We confirm that photon noise prevails over the background one for the Thomson data from the 2014 KSTAR campaign.
Figure 1. Measured filter functions (‘+’), i.e., relative transmittance coefficients of the optical band-pass filters as a function of the wavelength, for five separate channels of a core polychromator. Channel 1 looks at the shortest wavelength range, while channel 5 contains the Nd-YAG laser wavelength of 1064 nm. Also shown (lines) are the spectral distributions of the scattered photons [5] for various temperatures.

2 Estimating electron temperature

A polychromator contains five channels consisting of five band-pass filters shown in figure 1 with five photon detectors (IR-enhanced Si APD: Hamamatsu S11519). Figure 2 shows an example of the Thomson signal from each channel. Although the background noise levels are quite similar for all five channels, ch 1 does not show observable Thomson scattered signal. This may be due to electron temperature being less than 1 keV (see figure 1). The signal from ch 5 is mostly due to the straylight as attested by the fact that the signal levels before and during plasma shots are identical (not shown in the figure). The level of Thomson signal from ch 4 is smaller compared to signals obtained from chs 2 and 3 as the expected electron temperature is higher than 500 eV in the core of KSTAR plasmas. To estimate the integral of the $i^{th}$ channel Thomson signal $V_{TS}^{i}$, we perform a Gaussian fitting to the Thomson signal $V_{TS}$ (red line in figure 2(b)) which can be written as [4]

$$V_{TS}^{i} (t) = G n_{e} m_{laser} (t) \frac{d\sigma_{TS}}{d\Omega} \Delta Q L T(\lambda_{L}) QE \int d\lambda \frac{\phi^{i}(\lambda)}{\phi(\lambda_{L})} S(\lambda, T_{e}, \theta) \lambda_{L}$$

where $G$ is the APD gain factor, $m_{laser} (t)$ the number of photons per unit time as a function of time where $\int dt m_{laser} (t) = N_{laser}$ is the total number of photons in a single laser pulse. $\frac{d\sigma_{TS}}{d\Omega}$ is the differential Thomson scattering cross-section area, $\Delta Q$ the solid angle of the TS system and $L$ the scattering length. The net transmission coefficient of the collective optics $T(\lambda)$ is a function of the wavelength, but we assume that it is constant within the range of interests; hence $T(\lambda) = T(\lambda_{L})$ where $\lambda_{L}$ is the laser wavelength. As the $i^{th}$ channel filter function $\phi^{i}(\lambda)$ includes the wavelength variation of the quantum efficiency $QE$ of the APD detector, the value of $QE$ is taken at the laser wavelength. $S(\lambda, T_{e}, \theta)$ is the spectral distribution of the scattered photons [5] where $\theta$ is the
Figure 2. Measured Thomson signal from each channel for KSTAR shot #10433 at plasma time $= 10s$. Thomson signal from ch 1 cannot be distinguished from the noise, while the signal from ch 5 is dominated by the straylight. Red line in (b) illustrates the Gaussian fitting with the width of $w$ to the raw Thomson signal (black). Blue shaded region is used to estimate the background noise level of ch 2.

scattering angle. Then, $A_{TS}^i$ is simply

$$A_{TS}^i = \int dt V_{TS}^i (t). \quad (2.2)$$

From eq. (2.1), we find that the terms outside the integral do not depend on the channels. Thus, by taking the ratio of measured signals between the two channels such as

$$\frac{A_{TS}^i}{A_{TS}^j} = \int d\lambda \frac{\phi^i (\lambda)}{\phi^j (\lambda)} S(\lambda, T_e, \theta) \equiv R^{ij} (T_e, \theta), \quad (2.3)$$

we can construct an equation that depends on only $T_e$ and $\theta$. Note that we have $\theta \approx 90^\circ$ in this study [6]. Defining $M^i (T_e) = \int d\lambda \phi^i (\lambda) S(\lambda, T_e, \theta = 90^\circ)$, figure. 3(a) shows $M^i$ as a function of $T_e$ for all five channels. As we decide not to use signals from chs 1 and 5 due to indiscernible Thomson signals from the background noise or stray light, we create $R^{ij}$ with chs 2, 3 and 4 which are shown in figure. 3(b).
Figure 3. (a) $M^i (T_e)$ and (b) $R^{ij} (T_e)$ the ratio of $A_{iTS}^i$ between channels of 2, 3 and 4.

Because we may find three different $T_e$'s based on the measurements from figure 3(b), the $T_e$'s can be averaged with a weighting factor of the inverse of corresponding total uncertainty as

$$T_e = \sum_{i,j=2}^{4} T_{ij}^e \frac{1}{\sigma_{ij}^{total}} \left( \sum_{m,n=2}^{4} \frac{1}{\sigma_{mn}^{total}} \right)^{-1}, \quad (j > i, n > m),$$

(2.4)

where $T_{ij}^e$ and $\sigma_{ij}^{total}$ are the estimated electron temperature using $R^{ij}$ and its total uncertainty, respectively. $\sigma_{ij}^{total}$ can be estimated using a propagation of uncertainty assuming that the total uncertainty of the $i$th and $j$th channels, $\sigma_{total}^i$ and $\sigma_{total}^j$, are uncorrelated. Ideally, $\sigma_{total}^i$ needs to include both the background noise $\sigma_{bg}^i$ and the photon noise $\sigma_{ph}^i$:

$$\left( \sigma_{total}^i \right)^2 = \left( \sigma_{bg}^i \right)^2 + \left( \sigma_{ph}^i \right)^2. \quad (2.5)$$

However, not knowing the photon noise level, we estimate the $T_e$ by setting $\sigma_{total}^i = \sigma_{bg}^i$. In this paper, we estimate the background noise of the $i$th channel as

$$\sigma_{bg}^i = \delta^i \sqrt{N_{TS}^i \Delta t}, \quad (2.6)$$

where $\delta^i$ is the standard deviation of the data in the blue shaded region shown in figure 2(b). Due to the existence of Q-switch noise before the Thomson scattered signal which slowly decays away, we only use the blue shaded region, i.e., after the Thomson signal, to evaluate the background noise level. As shown in figure 2(b), we first fit a Gaussian function to the Thomson signal which finds the amplitude, peak-time and width ($w$) of the Gaussian. The blue shaded region where we obtain the background noise level starts at $5w$ away from the peak-time assuming that no Thomson signal exists in this region, hence only the background signal. $N_{TS}^i$ is the number of data points within the fitted Gaussian, i.e., $N_{TS}^i = 4w/\Delta t$, where $\Delta t$ is 0.8 ns, the time step of the data points with 1.25 GS/s. figure 4 shows examples of the temporal evolution of $T_e$ for two different KSTAR plasma shots estimated using eq. (2.4).
3 Synthetic Thomson data and signal-to-noise ratio

As the estimated $T_e$ shows about 15% (on average) variation on $T_e$ with respect to its mean value (e.g. see figure. 4), we raise the following questions: 1) can the background noise explain such an observed variation on $T_e$? 2) if not, how large does $\sigma_{ph}^i$ have to be to explain the observation? To be able to answer these questions, we generate synthetic Thomson signal where we can vary the noise levels numerically.

By letting $n_{laser}(t)$ in eq. (2.1) have a Gaussian form in a time domain (consistent with fitting the Gaussian function to measured Thomson signals), we generate the synthetic Thomson data for each channel similar to the measured data shown in figure. 2. We set the time step of synthetic data to be 0.8 ns corresponding to 1.25 GS/s of the fast sampling digitizer. A random number selected from a normal distribution is, then, added to each time point as a noise. The width of this normal distribution $\delta_{SYN}$ is set such that $\sigma_{SYN} = \delta_{SYN}\sqrt{N_{TS}\Delta t}$ (cf. eq. (2.6)) controlling the noise level. This completes generating synthetic Thomson signal for a single laser pulse. Let us denote the $i$th channel synthetic Thomson signal as $V^i_{SYN}(t)$ and its integral as $A^i_{SYN}$ estimated by fitting a Gaussian function to the $V^i_{SYN}(t)$ as if it were real measured signal.

We define the signal-to-noise ratio (SNR) of $T_e$, $\text{SNR}_{T_e}$, and that of the measured (synthetic) Thomson signal from the $i$th channel, $\text{SNR}^i_T$ ($\text{SNR}^i_{SYN}$), to be

$$\text{SNR}_{T_e} = \frac{\langle T_e \rangle}{\sigma_{T_e}}, \quad \text{SNR}^i_T = \frac{\langle A^i_{TS} \rangle}{\langle \sigma^i_T \rangle}, \quad \text{SNR}^i_{SYN} = \frac{\langle A^i_{SYN} \rangle}{\langle \sigma^i_{SYN} \rangle},$$

(3.1)

where the subscript * takes any one of the total (‘total’), background (‘bg’) or photon (‘ph’) for the noise source, i.e., $\text{SNR}^i_{bg}$ indicates the SNR of the Thomson signal due to the background noise. Here, $\langle \cdot \rangle$ indicates the time average with 50 consecutive data points equivalent to the time duration of 0.5 seconds, a couple of equilibrium evolution time scale for a typical KSTAR plasma, with the laser pulse repetition rate of 100 Hz during the 2014 KSTAR campaign. $\sigma_{T_e}$ is the standard
deviation of $T_e$ during the 0.5 seconds. Figure 5 shows time evolution of $\text{SNR}_{T_e}$ and $\text{SNR}_{i \text{bg}}$ for $i = 2$, 3 and 4 for the plasma shots shown in figure 4.

4 Estimating photon noise level and photon counts

We wish to find $\text{SNR}_{i \text{total}}$ and the corresponding $\sigma_{i \text{total}}$ for $i = 2, 3$ and 4 consistent with the experimentally measured $\text{SNR}_{T_e}$. Once we have $\sigma_{i \text{total}}$, $\sigma_{i \text{ph}}$ can be calculated using eq. (2.5) with the measured $\sigma_{i \text{bg}}$. Thus, we obtain $\text{SNR}_{i \text{total}}$ and $\sigma_{i \text{total}}$ with the synthetic data by recognizing that $\text{SNR}_{i \text{total}}^{\text{SYN}} = \text{SNR}_{i \text{bg}}^{\text{SYN}}$ and $\sigma_{i \text{total}}^{\text{SYN}} = A_{T_e} \text{SNR}_{i \text{total}}^{\text{SYN}}$. Here, we face a problem: there may exist infinite number of solutions on the combination of $\text{SNR}_{i \text{total}}$ for $i = 2, 3$ and 4 consistent with the measured $\text{SNR}_{T_e}$. To circumvent such a problem, anticipated by the experimental observation, we generate synthetic data only for chs 2 and 3 and let $\text{SNR}_{i \text{total}}^{\text{SYN}} = \text{SNR}_{i \text{bg}}^{\text{SYN}}$ since $\text{SNR}_{i \text{bg}}^{\text{SYN}} < \text{SNR}_{i \text{bg}}^{\text{SYN}} \sim \text{SNR}_{i \text{bg}}^{\text{SYN}}$ (see figure 5). We generate a database of $\text{SNR}_{T_e} = f \left( \text{SNR}_{i \text{SYN}}^{\text{SYN}}, T_e \right)$ for many different values of $\langle T_e \rangle$ for $i = 2$ and 3. Figure 6 shows the level of $\text{SNR}_{T_e}$ estimated with the synthetic data as a function of $\text{SNR}_{i \text{SYN}}^{\text{SYN}}$ at $\langle T_e \rangle = 800$ eV.

As the $\text{SNR}_{T_e}$ from the synthetic data has been estimated with only chs 2 and 3 (e.g. figure 6), we recalculate the experimental $\text{SNR}_{T_e}$ using only these two channels as shown in figure 7. Here, we also plot the ‘required’ level of $\text{SNR}_{i \text{total}}^{\text{SYN}}$ (green square) using the database of $\text{SNR}_{T_e} = f \left( \text{SNR}_{i \text{SYN}}^{\text{SYN}}, T_e \right)$ which allows us to estimate $\sigma_{i \text{total}}^{\text{SYN}}$ for $i = 2$ and 3 using eq. (3.1).

Figure 8 shows, for the KSTAR #10433, the levels of total required noise obtained by the synthetic data, background noise from experimental data and photon noise with eq. (2.5) normalized to the experimental signal levels for chs 2 and 3. It is clear that the KSTAR Thomson scattering system during 2014 campaign is photon noise dominated at least for the chs 2 and 3 of the polychromator we have investigated. This is good that the system is not limited by the background noise, however it also suggests that the system needs to collect more photons and utilize the collected photons more efficiently to reduce the 15% variation on the measured $T_e$. 

---

**Figure 5.** Examples of $\text{SNR}_{T_e}$ (black circle) and $\text{SNR}_{i \text{bg}}$ for $i = 2$ (red asterisk), 3 (blue diamond) and 4 (green triangle) for (a) KSTAR shot #10401 and (b) #10433. Here, $\langle T_e \rangle$ is estimated using chs 2, 3, and 4.
Figure 6. With synthetic data of chs 2 and 3, $SNR_{Te}$ is estimated as a function of $SNR_{SYN}$ at $\langle T_e \rangle = 800$ eV. This allows us to obtain the required values of $SNR_{SYN}^i$ and the corresponding $\sigma_{SYN}^i$ given the experimental $SNR_{Te}$.

Figure 7. Experimentally measured $SNR_{Te}$ (black circle) and $SNR_{bg}^i$ for $i = 2$ (red asterisk) and 3 (blue diamond) for (a) KSTAR shot #10401 and (b) #10433. Green squares show the ‘required’ $SNR_{total}^i$ estimated with the synthetic data consistent with the observed $SNR_{Te}$. Here, $\langle T_e \rangle$ is estimated using chs 2 and 3 only.

From the estimated normalized photon noise level, i.e., $1/SNR_{ph}^i$, we can estimate the ‘effective’ photon counts $N_{ph}^{i, eff}$ detected by the photon detectors as

$$\frac{1}{SNR_{ph}^i} = \frac{1}{\sqrt{N_{ph}^{i, eff}}} = \sqrt[4]{\frac{F}{N_{ph}^i}},$$

where $F$ is the noise figure describing the excess noise in the APD [4] provided that the APD shot...
Figure 8. Total (black circle), background (red asterisk) and photon (blue diamond) noise levels normalized to the signal levels of (a) ch 2 and (b) ch 3 for KSTAR plasma shot #10433.

noise is much larger than the total dark current noise, and $N_{ph}^i$ is the ‘actual’ photon counts. A calibration with the fast digitizers would provide a precise value of the noise figure. Since this calibration is not available the effective (rather than the actual) number of photons is estimated. Using the values from figure. 8 and eq. (4.1), we find that the effective photon counts are approximately 170 for chs 2 and 3. Note that the noise figure $F$ is always greater than or equal to 1, thus actual photon counts may well be larger than 170, i.e., by a factor of $F$. However, this cannot be regarded as the gain in photon counts since $F$ decreases the SNR.

The core KSTAR 2014 TS system has an effective F/# of 6.7 which corresponds to 17.5 msr. With the incident laser wavelength of 1064 nm and energy of 2.0 J, the incident photon number is estimated to be $1.1 \times 10^{19}$. Having approximate values of effective quantum efficiency of 15% (quantum efficiency of APD divided by the noise figure $F$), scattering length of 10 mm, optical transmittance of 40%, filter transmittance of 70%, we estimate effective photon budget to be approximately 6,200 photons per $1.0 \times 10^{19} \text{ m}^{-3}$ of electron density. For $T_e = 800$ eV, we find that the optical band-pass filters for chs 2 and 3 covers about 30% of the full spectrum (see figure. 1). Thus, the expected effective photon counts on chs 2 and 3 are approximately 1,800 about an order of magnitude higher than the estimated effective photon counts. This means that the KSTAR TS system has large potential to be improved.

5 Conclusion

In this paper, we estimated the time evolution of electron temperature by using the weighted look-up table method and found a 15% variation on $T_e$. By using the synthetic Thomson data, we have found that such a variation on electron temperature cannot be explained solely by the background noise and thus estimated the required uncertainty level consistent with the experimental observation. Our study indicates that the chs 2 and 3 of the polychromator we have investigated are dominated by
the photon noise with the average effective photon counts per laser pulse of 170. This suggests that improving the photon collection system and decreasing the noise figure of the system is required to decrease the observed scatters in electron temperature.

Acknowledgments

Authors would like to thank the referee for his/her valuable comments. This work is supported by National R&D Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (grant number 2014M1A7A1A01029835) and the KUSTAR-KAIST Institute, KAIST, Korea.

References

[1] S.T. Oh, J.H. Lee and H.M. Wi, Examinations of electron temperature calculation methods in Thomson scattering diagnostics, Rev. Sci. Instrum. 83 (2012) 10D525.

[2] J.H. Lee, S.T. Oh and H.M. Wi, Development of KSTAR Thomson scattering system, Rev. Sci. Instrum. 81 (2010) 10D528.

[3] B.P. LeBlanc, Thomson scattering density calibration by Rayleigh and rotational Raman scattering on NSTX, Rev. Sci. Instrum. 79 (2008) 10E737.

[4] R. Scannell, Investigation of H-mode edge profile behaviour on MAST using Thomson scattering, Ph.D. Thesis, Department of Electrical and Electronic Engineering, UCC, NUI (2007).

[5] O. Naito, H. Yoshida and T. Matoba, Analytic formula for fully relativistic Thomson scattering spectrum, Phys. Fluids B 5 (1993) 4256.

[6] S. Oh and J.H. Lee, Examination of a duo-collection optics design for the Korea superconducting tokamak advanced research (KSTAR) Thomson scattering system, Meas. Sci. Technol. 22 (2011) 035303.