The 8th International Conference on Optical Terahertz Science and Technology, paper 74, We-P-46, USA, (2019)

Statistical denoising models for THz time-domain spectroscopy

Maksim Skorobogatiy, Jayesh Sadasivan, Kathirvel Nallappan, Hichem Guerboukha

Engineering physics, Polytechnique de Montréal, C. P. succ. Centre-ville, Montréal (QC), Canada H3C 3A7

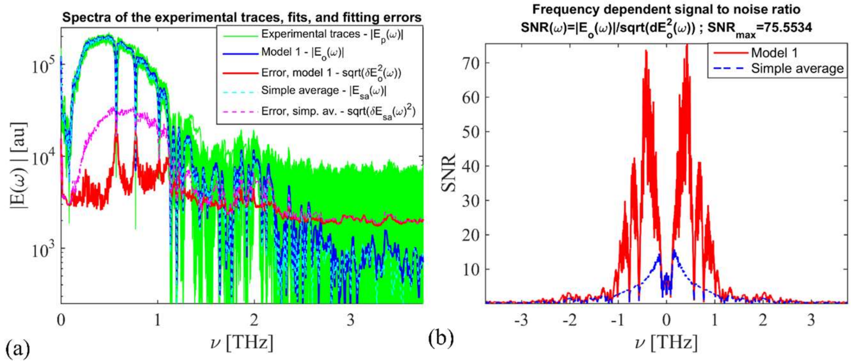
As a general class of optical experiment, the pump-probe technique is used in many different physical and engineering fields. Within this technique, a laser beam (pump) is used to excite a sample, while a time delayed laser (probe) interacts with the excited sample to provide temporally precise dynamics of a physical phenomenon. In general, while all the measured pulses contain identic physical information, each one of them also includes a unique noise contribution. The goal of this work [1] is to develop generic mathematical models to extract the information and mitigate noise. In some regards, the THz time-domain spectroscopy (THz-TDS) can be seen as a pump-probe experiment. Therefore, we also validate our models using experimental data obtained with a THz-TDS.

The most straightforward way of treating a collection of pump-probe pulses is to take a simple average of the pulse $E_v(t)$: $E_{sa}(\omega) = 1/N \sum_{v=1}^{N} E_v(t)$. This model assumes that the noise $\delta E_v(t)$ is linearly added to the nominal signal, while in reality there are several types of noise in a pump-probe experiment that do not necessarily contribute in a linear additive fashion. In this work, we developed 3 models for averaging of the pump-probe pulses:

Model 1 $E_v(t) = C_v E_0\big(t - \delta t_v\big) + \delta E_v(t - \delta t_v)$
Model 2 $E_v(t) = C_v \big|E_0\big(t - \delta t_v\big) + \delta E_v(t - \delta t_v)\big|$
Model 3 $E_v(t) = C_v E_0\big(t - \delta t_v\big)\big|1 + \delta_v\big(t - \delta t_v\big)\big|$

In the above expression, C_v is the complex gain factor that account for “slow” noise (laser power variation, emitter antenna aging, etc.), δt_v is the temporal shift from one measurement to another and $\delta E_v(t)$ is the leftover “fast” noise that cannot be compensated using a lock-in amplifier for example. In Model 1, the “fast” noise is independent of the pulse amplitude C_v, which is a valid assumption if the noise is purely electronic in nature. In Model 2, the “fast” noise is proportional to the amplitude of the excitation laser beam C_v. Finally, Model 3 assumes that the “fast” noise is incurred during pulse propagation due to rapidly changing environmental factors (air flows, sudden humidity variation etc.).

For each model, we develop algorithms [1] by solving a minimization problem of the noise. Next, we use 400 experimental traces obtained from a THz-TDS to confirm the validity of our models. In Fig. 1(a), we compare the fit using Model 1 to a standard average approximation. We observe that the error in the simple average is larger than the Model 1 error. This is further confirmed by computing the signal-to-noise ratio in Fig 1(b). The SNR of Model 1 reaches a maximum of 75, while the standard average gives a maximum SNR of 15.



[1] M. Skorobogatiy, J. Sadasivan, and H. Guerboukha, “Statistical models for averaging of the pump-probe traces: example of denoising in terahertz time-domain spectroscopy,” IEEE Trans. THz Sci. Technol. **8**(3), (2018).