Exploring modern alternatives to the Whittaker-Shannon-Nyquist sampling theorem in THz Spectroscopy

Christian D. Andersen[1‡], Jacob A. S. Noura[1‡], Bjørn H. Mølvig[1], Oscar G. Garcia[1], Simon J. Lange[1*]
[1]*Department of Electrical and Photonics Engineering, Technical University of Denmark, Kgs. Lyngby, Denmark*
**slla@fotonik.dtu.dk*
[‡] the two Co-Authors share equal contribution to the work

***Abstract*—We present simulations for a THz cross-correlation spectroscopy (THz-CCS) optical system. The aim is using compressed sensing (CS) to reconstruct the THz signal from a random under sampling of the signal and potentially replacing a delay stage unit of the THz-CCS system to increase robustness and cost-effectiveness of the optical system. We present results from using the CS to attempt to reduce the number of samples required to interpolate the THz signal in question. We compared Shannon-Whittaker interpolation to CS reconstruction in a multitude of basis. We conclude that within the current implementation of CS Shannon-Whittaker interpolation beats CS for our simulated THz signal. We have however only just begun to explore the multitude of potential CS implementations for THz spectroscopy systems and see a host of additional explorations to be conducted in the near future.**

I. Introduction

A key technological field for modern engineering and physics is emission and detection of electromagnetic (EM) radiation. Different frequency regimes necessarily need different devices for such task. Below 300 GHz devices based on alternating currents are used, but prove obsolete above such frequencies. Above 40 meV (10 THz) lasers and light diodes are common as emitter devices, furthermore photon energies in this regime are high enough for detected matter interaction [1]. But above 300 GHz and below 10 THz emission and detection setups are typically expensive and unfit for most practical uses. This technical challenge has historically limited the ability of THz technology to become widespread across multiple fields of science and industrial applications. Today, THz technoloy therefore holds a huge, untapped potential. One potential THz technology platform that holds great promise is THz cross-correlation spectroscopy (THz-CCS) [2].

In this work we use compressed sensing (CS) on a THz CCS optical system in order to explore the potential for optimizing the speed and flexibility of the system. The optical setup and a phenomenological illustration of the working principle for detection can be seen on fig. 1. The broadband (BB) source is split into an emitter arm and detector arm where a delay line probes the signal. The source is a BB continuous source, which constitutes a considerable cost-effectiveness improvement from other THz time-domain spectoscopy (THz-TDS) systems using femto-second mode locked lasers. The measured signal is sampled at the Whittaker-Shannon-Nyquist rate such that Shannon-Whittaker interpolation can be used to reconstruct the signal. The signal is furthermore Fourier transformed and features from the sample can be extracted with both phase information and spectral amplitude information. Furthermore we investigate schemes for undersampling the signal in pursuit of long term finding alternatives to the delay line scheme as such component either limits size and portability of the full THz-CCS system or reduces robustness of the system considerately, both being unfavourable for a broadly applicable system. We simulate the THz-CCS system with both single delay line as in fig. 1 and multiple delay lines as this seeks to potentially improve performance of CS signal reconstruction. A multitude of different potential sparse basis are examined for reconstruction of the signal with respect to the degree of undersampling. The results are compared to Shannon-Whittaker interpolation for the same degree of sampling.

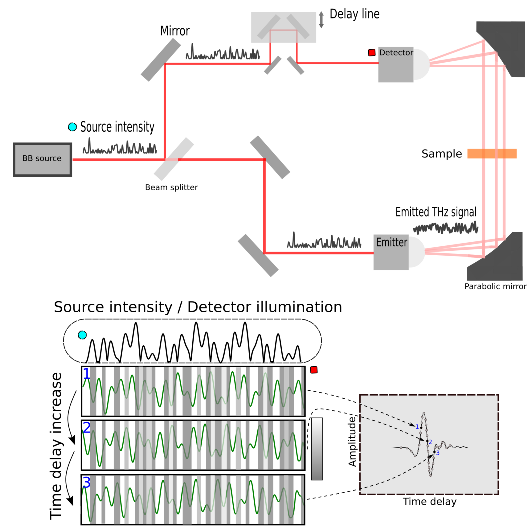

Fig. 1. Optical setup for a THz-CCS system with a phenomenological illustration of the working principle of THz-CCS. A broadband source is used to generate continuous wave incoherent light which is split into a detector arm with a delay line and an emitter arm with a photoconductive antenna (PCA) emitting continous THz light. The THz light is passed through a sample or air and the detector detects a THz time-signal as a function of the time-delay from the delay-stage

II. Results

In Fig. 1, a conventional CCS system is illustrated. This is modelled in a numerical simulation in Python, where a model of a C-band Amplified Spontaneous Emission (ASE) source, is used as the broadband source. The analytical expression for the resulting current in the antennas [4] is then used to create the resulting signal, which we choose to calculate as a 128-point vector, which we normalize, so the highest value of the vector is 1.

We now used compressed sensing, to recreate the signal. We choose to use the originally proposed Compressed Sensing model [3], instead of LASSO regression, as is commonly used. This is a more strict constraint. As we only have a single delay line, our measurement matrix has only a single non-zero digit per row, and that digit is of course 1. In order to maximize incoherence between the measurement and basis changing matrix, these 1's on each row are randomly placed. The discrete basis changing matrices, were chosen as the two most used in the current literature, which is the sine and cosine transforms, and a subset of the classical orthogonal polynomials (Chebyshev and Legendre). The Hermite, Laguerre and Canonical polynomial transforms, were not considered as these have numerical convergence issues. This is most likely due to them being too coherent, with our measurement matrices.

As each measurement matrix is random, the recreations vary, and the RMSE (Root-Mean-Square Error), between the recreation and the true signal, also varies. The signal is therefore recreated 100 times for each sample point, and the RMSE is calculated between each of these and the true signal, and the average RMSE is then calculated. This is pictured in Fig. 2 **(a)**. The RMSE of the Whittaker-Shannon-Nyquist interpolation (Sinc interpolation) and the true signal, is also pictured in Fig. 2 **(a)**, as a comparison. Note that none of the basis transforms is able to outperform, that is, reach zero RMSE for number of sampling points, before the Sinc interpolation. We note that the RMSE is initially lower for the compressed sensing, than for the Sinc interpolation. This is due to them being more noisy, and thereby being on average closer to zero.

Fig. 2 **(b)**-**(c)** respectively include the minimum amount and an insufficient amount of sample points, needed for the Sinc interpolation.

As each RMSE valued calculated for Fig. 2 **(a)**, is an average of 100 recreations, it is also possible to calculate the standard deviation associated with each RMSE. This is pictured in Fig. 2 **(d)**, for the discrete sine transform, together with the RMSE of the Sinc interpolation. We again do not beat out Whittaker-Shannon-Nyquist, even with the best random measurement matrices. We however note, that there indeed is a large difference between "lucky" and "unlucky" measurement matrices, leading to respectively "good" and "bad" recreations. This is an often overlooked fact in Compressed Sensing applications.

In conclusion, conventional compressed sensing basis, that has seen great results in other fields, does not beat out the

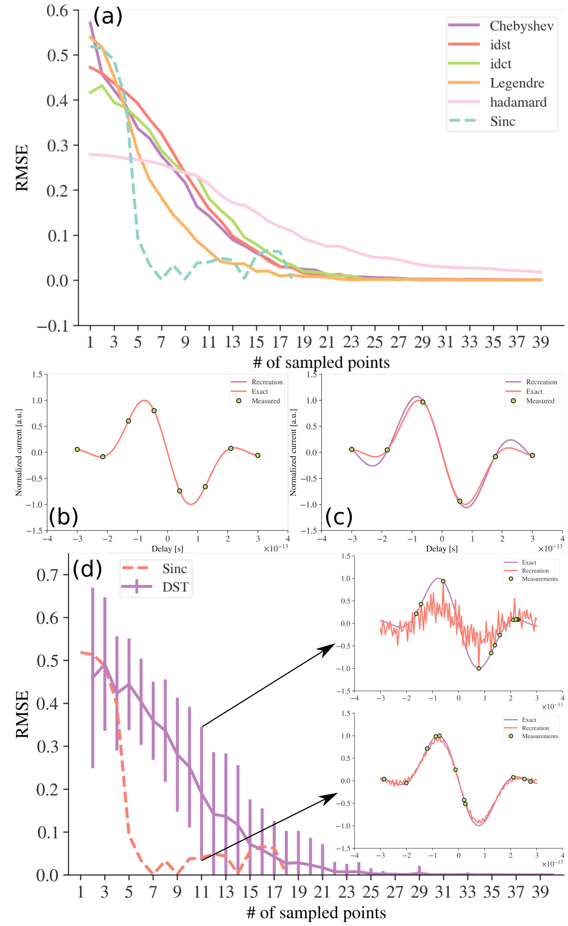

Fig. 2. **(a)** RMSE (Root-Mean-Square Error), of conventionally used basis in Compressed Sensing (CS), and RMSE for the Sinc interpolation, for a 128-point signal. As each measurement matrix for a simulation is random, the RMSE for each number of sample points is an average of 100 simulations. The RMSE for the Sinc interpolation becomes positive after reaching zero due to the numerical integer rounding of indices when taking evenly spaced points. **(b)** Sufficiently sampled Whittaker-Shannon-Nyquist (Sinc) interpolation. **(c)** Insufficiently sampled Whittaker-Shannon-Nyquist (Sinc) interpolation. Note that there is only one less sample point than in **(b)**. **(d)** RMSE of the Sine transform, and its associated standard deviation, as a result of the 100 simulations. Also pictured is the RMSE of the Sinc interpolation.

Whittaker-Shannon-Nyquist interpolation when considering THz Spectroscopy. However more advanced CS methods, such as utilizing specifically created basis, hold great promise for further exploration in the near future.

References

[1] Jepsen, P., Cooke, D. and Koch, M. (2011), Terahertz spectroscopy and imaging – Modern techniques and applications. Laser Photon. Rev., 5: 124-166. https://doi.org/10.1002/lpor.201000011

[2] Daniel Molter, Michael Kolano, and Georg von Freymann, "Terahertz cross-correlation spectroscopy driven by incoherent light from a superluminescent diode," Opt. Express 27, 12659-12665 (2019)

[3] Candès, E.J., Romberg, J.K. and Tao, T. (2006), Stable signal recovery from incomplete and inaccurate measurements. Comm. Pure Appl. Math., 59: 1207-1223. https://doi-org.proxy.findit.cvt.dk/10.1002/cpa.20124

[4] Maik Scheller and Martin Koch, "Terahertz quasi time domain spectroscopy," Opt. Express 17, 17723-17733 (2009)