Broadband super-resolution Terahertz Time domain spectroscopy applied to Gas analysis

Sophie Eliet, Arnaud Cuisset, Francis Hindle, Jean-Francois Lampin, Romain Peretti

Broadband Super-Resolution Terahertz Time Domain Spectroscopy Applied To Gas Analysis

Sophie Eliet, Arnaud Cuisset, Francis Hindle, Jean-François Lampin, and Romain Peretti

***Abstract*—Commercially available Terahertz (THz) Time domain spectroscopy (THz-TDS) apparatus are used in multiple application fields from material science to biology, and even in industry to measure the thickness of a paint layers. Its broad spectral coverage is of interest for air quality monitoring and atmospheric composition analysis. However, gas spectroscopy needs high resolution that the usual approach in THz-TDS is unable to provide. In this letter, we introduce the concept of constraint reconstruction for super-resolution spectroscopy based on the modeling of the spectroscopic rotational lines in a sparse spectrum. It is demonstrated for low-molecular weight molecules in gas phase that typically have spectra containing a small number of narrow lines over a broad spectrum. This alternative approach directly reconstructing the spectra achieved a resolution improvement factor of 10 compared to the best available resolution by the standard THz-TDS data treatment.**

***Index Terms*—Time domain spectroscopy, Terahertz gas spectroscopy, super resolution spectroscopy**

I. Introduction

Terahertz (THz) Time Domain Spectroscopy (THz-TDS) began its development with the birth of the femtosecond laser as sub-picosecond sampling [1]. The idea was quickly used for spectroscopy [2] and few years later, the integration of photoconductive antenna [3, 4] bring it to a practical broadband THz spectroscopic system. These systems are now commercial and offer more than one decade of bandwidth. They enabled myriads of applications from semiconductor spectroscopy [5, 6], 2D materials [7] or biological samples [8]. The application of THz-TDS to study gas phase samples is attractive as the broad coverage up to 6 THz should enable fast multi-species analysis. In fact, this has been tried as early as in 1989 [9]. Nevertheless, gas spectroscopy studies employing THz-TDS are uncommon as the length of the delay line limits the resolution. Indeed, the time domain signal is recorded over a short time window corresponding the mechanical translation of the delay line. The spectrum is obtained using a Discreet Fourier Transform (DFT) whose resolution is determined by the time window, a limitation that we will call the Fourier criteria here (*Fc*).

To optimize the molecular selectivity gas spectroscopy and sensing applications are generally undertaken at low pressures (<10 mbar) where the collisional broadening is low. In this case, the apparatus should be able to observe a small number of narrow rotational lines requiring high-resolution that can be obtained with narrowband techniques [10, 11]. The application of THz-TDS could be imagined by employing very long delay lines. This would significantly degrade the signal to noise ratio, the stability and the compactness of the systems. Other experimental approaches based on the use of optical frequency combs and so called Asynchronous Optical Sampling System (ASOPS) [12] have overcome the limit of the delay stage and enabled new broadband spectroscopic setups of particular relevance for THz molecular spectroscopy. These systems are still, extremely complex to operate, costly and limited in bandwidth and signal to noise ratio compare to THz-TDS. Improvements were proposed [13] leading to better performances but requiring very time-consuming data accumulation periods.

Signal processing is nowadays used everywhere on a daily basis on myriads of applications from the most basics smartphone camera using data treatment to improve image quality to advanced one in military radars enabling the detection of very small objects. Therefore, advanced signal processing techniques have already been applied in many fields: quantum chemistry [14], nuclear magnetic resonance [15] and photonic simulations [16]. Methods such as wavelet transform [17], filter diagonalization [18] or harmonic inversion [14] have allowed resolution performance improvements far beyond *Fc* because they do not use basis functions without pre-fixed central frequency as it is done in DFT. Super-resolution spectroscopy has already been demonstrated for Fourier Transform Interferometers (FT-IR) using maximum entropy methods [19] and compact representation [20].

Here we propose a Constrained Reconstruction Super Resolution (CRSR) algorithm for super-resolution spectroscopy from a commercially available broadband (~0.2 – 6 THz) THz Time Domain Spectrometer with a Fourier transform limited spectral resolution of *Fc* = 1.2 GHz. This algorithm is also applicable to other spectroscopic techniques using a frequency comb. We validated the methods on rotational spectra of gas phase ammonia (NH_3) with the measurements of Doppler limited and self-broadened rotational lines in the 1 – 700 mbar pressure range.

II. Constraint reconstruction super resolution method

A major difference between spectral resolution in spectroscopy and spatial resolution in microscopy is the shape of the detected object. Compared to a microscope, which targets an object with an arbitrary shape, a spectrometer measures lines following specific profiles constrained by the laws of quantum

mechanics. This "*a priori*" knowledge is very important since it allows a constrained reconstruction approach [21] without any information loss. In other words, in the frequency domain, rovibrational lines of molecules diluted in a gas sample follow in the simplest models a Gaussian, a Lorentzian or a Voigt profile, corresponding respectively to the Doppler, the collisional or a mix between these two types of broadening. At the pressures considered in this study (1 < P < 700 mbar), the collisional effects dominate (3000> $\frac{\Delta\nu_{coll}}{\Delta\nu_{Doppler}} > 2$) and only the Lorentzian profile has to be considered.

In terms of photonics propagation, this is described by a complex (j is the square root of -1) refractive index ($\tilde{n}$) or a dielectric constant $\tilde{\varepsilon}()$ dependence following the equation[1]:

$$\tilde{\varepsilon}(\nu) = \tilde{n}^2(\nu) = \varepsilon_\infty + \sum_{k=1}^{k_{max}} \frac{\Delta\varepsilon_k \nu_{0,k}^2}{\nu_{0,k}^2 - \nu^2 + j\nu\Delta\nu_k}\,, \qquad (1)$$

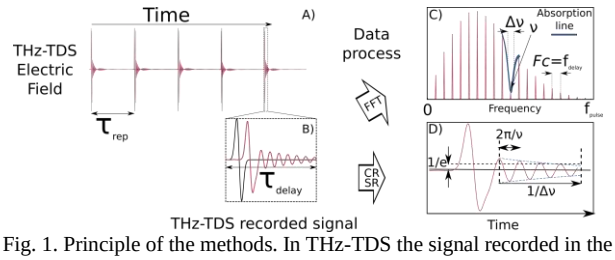

Fig. 1. Principle of the methods. In THz-TDS the signal recorded in the time domain (A) is limited by the time windowing coming from the travel range of the delay line from $t = 0$ to t=$t = \tau_{delay}$ (see zoomed part in B). Usually, a discreet Fourier transform is applied to this signal (C) and the spectral resolution is limited to $Fc = f_{delay} = {}^1/_{\tau_{delay}}$. The CRSR in (D) uses the same data and perform a full analysis of the time domain signal by fitting the parameters of damped sinus (bottom right) to retrieve the parameters of each spectral lines allowing the reconstruction of the super resolution spectrum.

To retrieve the three Lorentz parameters, we implemented the super-resolution in an open source software fit@TDS previously developed by our group [22, 23]. Here, we model the propagation of the THz wave in the experiments thanks to the complex transmission coefficient. We write it as the ratio of the signal with gas in the cell (sample) by the one without gas (reference).

$$\tilde{T}(\omega) = \frac{\tilde{E}_s(\omega)}{\tilde{E}_{ref}(\omega)} = exp\left(-j\frac{\omega l}{c}(\tilde{n}(\omega) - 1)\right), \qquad (2)$$

Here, $\omega = 2\pi\nu$, $\tilde{E}_s(\omega)$ and $\tilde{E}_{ref}(\omega)$ are the Fourier transforms of time-domain signals $E_s(t)$ and $E_{ref}(t)$, respectively, $\tilde{n}$ is the complex refractive index where the real part corresponds to a delay and the imaginary part to an absorption in the sample, l is the length of the gas cell. Here, thanks to the *Brewste*r angle design, we considered negligible the variation of the transmission term between the inner part of the cell and the window with and without gas.

Then we calculated $E_{model}\{p_i\}(t)$ (where p_i is the set of parameters for each oscillator defining the line center width and intensity) and t the time; simply by convoluting $\tilde{E}_{ref}(\omega)$ with $T\{p_i\}(t)$, combining equations 1 & 2. To quantify the validity of this model $E_{model}\{p_i\}(t)$, we compare it model with the measured sample time-trace using the following objective function to minimize:

$$Obj\{p_i\} = \sum_{t=0}^{t=t_{max}} \left(E_{model}\{p_i\}(t) - E_s(t)\right)^2 dt\,, \qquad (3)$$

Where the super–resolutions holds in t_{max} and will be discussed in the next paragraph.

It may be noted here, that we are going back and forth from time to frequency domains using the Fourier transform, as this mathematical concept was actually feasible. In reality, and as in most of experimental cases, Fourier transforms are not performable on our data because our data are sampled and bounded. As is generally employed, we are using the Fast Fourier Transform (FFT) to perform the time/frequency transformation. FFT is a DFT and, as such, differs from Fourier transform by two ways. It assumes the signal to be sampled and periodic. The super–resolution approach surpasses the previous version because it uses a different time frame (t_{max}) in the FFT. Previously and as it is widely done in the community, the time frame was set to be the length of the delay lines (τ_{delay} in fig 1.). This choice assumes that the signal is lower than the noise at the end of the time trace [24]. As shown in fig 1., in the case of spectral lines of width narrower than $1/\tau_{delay}$, the damped sinus corresponding to the Lorentz line spreads further than the delay line and thus does not fulfill the assumption. Precisely, if one would use τ_{delay} as t_{max} a line narrower than the *Fc* limitation will give rise to a temporal signal going over the right edge of the time window and coming back on the beginning of it on the left edge as explained in [13].

Here, we take the benefit of THz-TDS system to be time periodic and achieve super–resolution. In fact, the excitation

where ε_∞ is the dielectric permittivity at the highest frequency according to the range of interest, k_{max} is the number of considered oscillators, ν_{0k} , $\Delta\nu_{\,k}$ and $\Delta\varepsilon_k$ are the resonant frequency, the damping rate and the strength (expressed in permittivity units) of the k^{th} oscillator, respectively.

A Lorentzian profile corresponds, in the time domain, to an exponentially damped sinus function. The super-resolution relies on the fact that each damped sinus / oscillator is described by only three parameters: the amplitude, the central frequency and the damping rate. As seen on figure 1, only few time domain experimental points are required to retrieve the parameters. Still, such methods do not create information and the *Fc* limitation translates into the generalized *Fourier-Heisenberg* uncertainty that limits the number of oscillators to be fewer than a fraction of the total number of points in measured THz-TDS trace [14]. With this constraint, the resolution will only be limited by the experimental signal-to-noise ratio and the computational methods to optimize the fit.

[1] We took as a convention that all temporal functions are denoted with a letter and their Fourier transform with the same letter and a tilde.

laser has a repetition time τ_{rep} (fig. 1) that we used as t_{max}. Therefore, the FFT follows the same periodicity than the experiments preventing any artificial DFT-folding artifact or aliasing and even reproducing the ones from the experiments. Certainly, since the length of the delay line does not allow the recording of the full time-frame. Consequently, we fill the remaining data for the reference measurements (without gas, so without Lorentz line narrower than $1/\tau_{delay}$). The only assumption made here is that the emission from the photoconductive antenna is below the noise after τ_{delay}. Concretely, we zero feed $E_{ref}(t)$ for every t between τ_{delay} and τ_{rep} to build $E_{model}(t)$ using equations (1 & 2). Then, we compared only the recorded fraction to the model while keeping t_{max}= τ_{delay} in equation (3) because $E_{sample}(t)$ is not recorded for t between τ_{delay} and τ_{rep}. We used this residual error as the value to minimize in the CRSR. This is an optimization problem set in either an augmented Lagrangian particle swarm [25] or a sequential least squares [26] optimizer. The result of this optimization is the set of parameters defining the k oscillators.

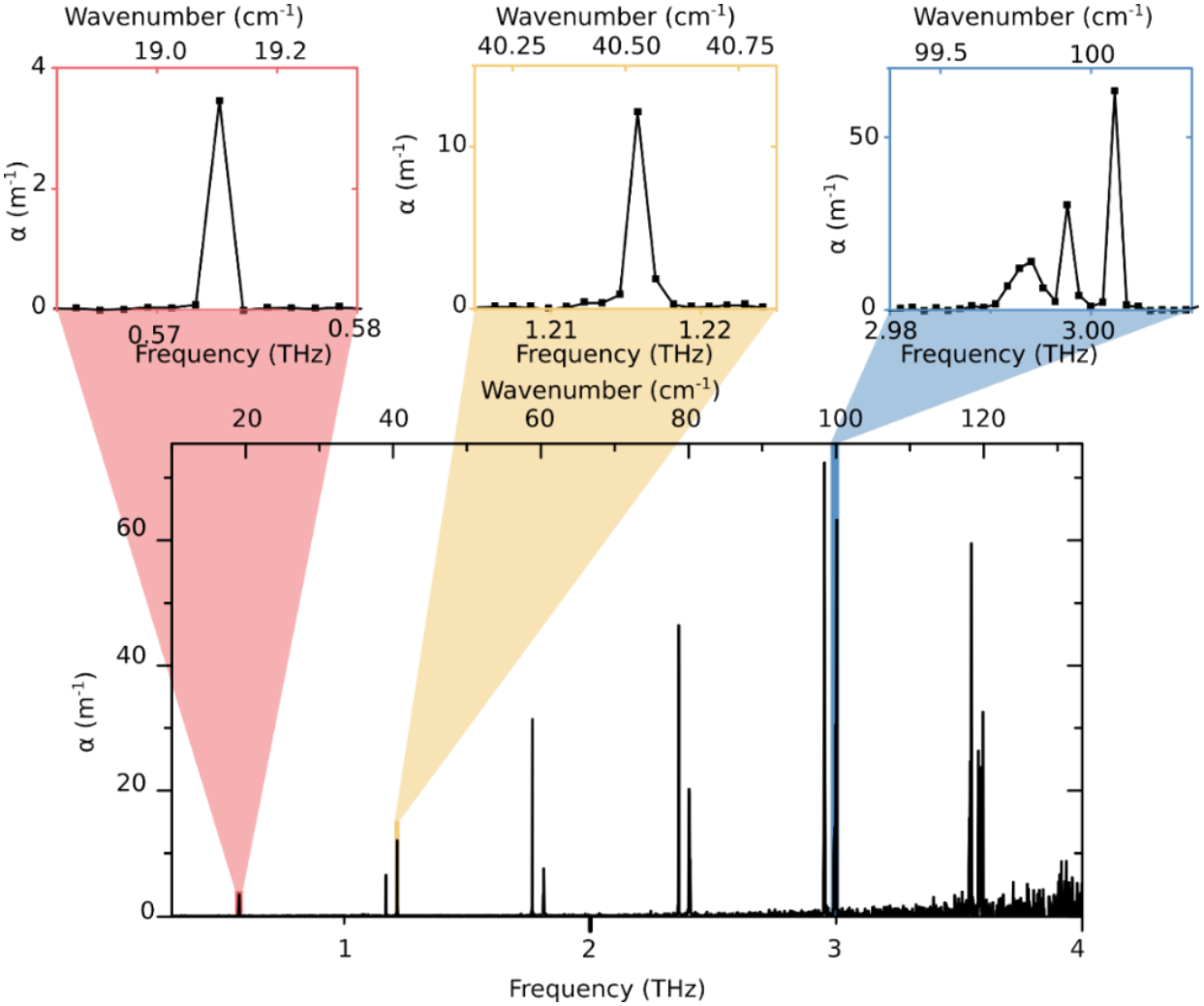

Figure 2: Absorbance THz-TDS spectrum of gaseous NH_3 at 10 mbar (simple FFT) showing that the spectrum extends from 200 GHz to 4 THz with narrow lines distributed from 500 GHz to 3.6 THz. It was recorded using a 7 cm long Brewster angle gas cell. The three top insets zoom the lines specifically studied in the paper (see Fig. 3 and Fig. S4)

III. Experimental demonstration

We assessed the CRSR performances by carrying out absorption measurements of NH_3 using a commercial transportable all-fiber (except the delay line) THz-TDS system Terasmart from Menlo Systems GmbH. The THz beam was routed from and to the photoconductive antenna, using 50 mm focal length TPX lenses. To suppress the Fabry-Perot echoes we set a Brewster-angle High Resistivity Silicon-windows gas-cell, with an optical path length of 7 cm, in the collimated THz beam. The cell was evacuated with a turbomolecular pump and pure NH_3 (99.98%) was injected in the cell to obtain the targeted pressure. From 1 to 200 mbar a thermostated Baratron® capacitive gauge was used with a precision of 0.12%. From 200 mbar to atmospheric pressure, an Edwards CG 16K capsule dial gauges barometrically compensated was used with a precision of 3 %. We repeated the thirteen minutes long experiments (1000 acquisitions of 780 ms) for pressures from 1 mbar to 690 mbar. The time window was τ_{delay} = 848 ps wide (time sampling of 33 fs) leading to a *Fc* of 1.2 GHz, the

repetition frequency was f_{rep} = 100 MHz (τ_{rep} = 10 ns ~ 12 τ_{delay}).

To give a better insight on the studied spectrum, Figure 2 shows the recorded experimental spectrum of 10 mbar NH_3 gas with the THz-TDS over the whole bandwidth using FFT on the time trace. Here one sees the NH_3 rotational lines spreading over the broad THz-TDS spectrum up to 4 THz. It is clear here that the Fc limitation does not allow to resolve individual ammonia rotational line and to describe correctly the line profiles. Even worst doublets of lines separated by few MHz as expected in the 1.215 THz region are not visible. We applied the CRSR on the recorded time traces in a large range of pressures with k_{max} = 36 oscillators corresponding to 36 rotational lines observed from 200 GHz to 4 THz with tabulated line intensities between 1.10^{-20} and $4.6\ 10^{-19}$ cm^{-1}/molecule.cm^{-2}. In the supplementary document the table S1details for the 36 oscillators the associated rotational transitions and compares the CRSR retrieved parameters (frequencies and linewidths) with the literature values. The assessment of the CRSR performances may be performed regarding three key performances. First its ability to point the center of a line. Second its ability to determine its linewidth. Finally, its ability to distinguish two close spectral lines. The results are depicted in Fig.3.

Panel A shows the reconstructed spectrum of the isolated absorption rotational line around 572.5 GHz for three pressures (6, 55 & 580 mbar) compared to the FFT of the time trace. For the widest line measured at P=580 mbar, the CRSR gives similar results to the FFT. With the usual FFT method, at P=55 mbar and P=6 mbar, we can only state a line with a frequency centered at 572.5 ± 1.2 GHz and a linewidth below F_c. In the case of the minimization by CRSR gives centers of 572.52 and 572.72 GHz, and linewidths of 1 118 MHz and 114 MHz respectively for 55 mbar and 6 mbar. To evaluate the accuracy of this reconstruction, we plotted on panel B and C the variation of the extracted line-frequency and line-width with the gas pressure. These are compared the line parameters for the self-shift and the self-broadening coefficients tabulated in the HITRAN 2017 atmospheric database [27] using as reference the measurements performed by Belov *et al.* with high-resolution techniques [28]. The comparison demonstrates that the CRSR frequencies are determined with a precision of 20 MHz down to a pressure of 3 mbar, *i.e.* a relative precision of $3.5\ 10^{-5}$ at 573 GHz. This precision exceeds the one of the commercial continuous wave systems [29] and is even a measurement of the absolute central frequency thanks to the calibration done on water vapor lines (see S.I part IV). We attribute the major contribution to the line frequency error to the pressure measurement uncertainties for the highest pressures and to the decreasing of the SNR at lowest pressures.

The second criteria: the linewidth corresponds to the one of the database better than 5 % from atmospheric pressure down to 10 mbar before a progressive increasing of the error at lower pressures. We retrieved an accurate linewidth down to a pressure of 6 mbar corresponding to a FWHM of 114 MHz with a precision of 15 MHz, representing a resolution improvement close to two orders of magnitude compared to the *Fc* limitation.

To reach the third criteria and achieve the doublet separation, we focused on how to discriminate two absorption lines close together, namely the 1.21 THz $2_{0,(a)} \leftarrow 1_{0,(s)}$ and $2_{1,(a)} \leftarrow 1_{1,(s)}$ NH_3 rotational line doublet separated by ~393 MHz. Here we want to retrieve the two sets of parameters for each component of the doublet. Nevertheless, the strong correlation between these parameters, forced us to add an additional constraint to lift this complication. For the first and only time of the study, we fixed the ratio of oscillator strengths between the two lines to the database value of 0.38 [27]. The Panel D of figure 2 shows the doublet at 55 mbar. One can see here that even at higher pressure, when it is not possible to spectrally distinguish the two fully overlapped individual lines, the CRSR is able to retrieve the two sets of parameters for each line plotted on panels E and F. There, one can note the very good agreement with the tabulated values. At 24 mbar when the shape of the line begins to show a shoulder, and at 7.5 mbar when the two lines are separated, the retrieval is always very good. Once again, to check the reliability of the CRSR method, we compare the retrieved parameters to the HITRAN tabulated ones used as reference. On panel E, from the comparison of the central frequencies of both lines of the doublet at very low pressure, we estimate the absolute pointing precision to ~50 MHz, namely a relative precision of 4.10^{-5} at 1.215 THz. Moreover, on panel E, the comparison shows two limits on the line frequencies determination: at very low pressure (< 7 mbar) the SNR and at high pressure (> 80 mbar) the overlapping of the lines as demonstrated by the numerical example in SI (fig S1 and 1§). Still, the values are very accurate from a pressure of 7.5 mbar up to 55 mbar. In this range, the relative accuracy is better than 30 % for a line as narrow as 150 MHz. Finally, we can notice that, as shown by the numerical example given in supplementary material (fig S1) with a more favorable signal to noise conditions, the doublet may be retrieved precisely without constraining the ratio of the two oscillator strengths. This demonstrates that CRSR fulfils the third criteria for super resolution without any intrinsic limitations: we have clearly succeeded to reconstruct the two components of ammonia doublet separated by less than the usual resolution of the apparatus. It is important to comment that if there is no theoretical limitation for the resolution of such a method [14], the signal to noise ratio is the only limiting factor. A corollary of this assertion is that reducing the absorption of the pure gas sample by reducing the pressure will decrease the resolution. In addition, since the excitation source is a periodically pulse laser so-called laser frequency comb, the excitation THz signal is a comb itself. It means that the excitation spectrum is very similar to an enveloped *Dirac* comb. One of the consequences is that a line much narrower than f_{rep} will have only a small probability to absorb the incident comb spectrum and this will give an almost zero absorption signal. Still, this does not set a rigorous resolution limitation to f_{rep} as shown in fig. 3 and S1. Certainly, the very high dynamic range of THz-TDS setup allows retrieving the parameters of the lines thanks to the absorption from its shoulders, even if the center of the line is not tuned to one tooth of the comb. Therefore, if f_{rep} is not an absolute resolution limitation it is quite clear that measuring the

parameter of a line much narrower than f_{rep} will require a tremendously high (absorption)-signal to noise ratio limiting *de facto* the resolution to a fraction of f_{rep}.. This limitation stands because a stable frequency comb is used (f_{rep} variation of ~30 Hz). The use of a less stable comb (~ 20 kHz) will lead to the complete coverage of the free spectral range at 500 GHz. Another way of mitigation would be to modulate the f_{rep} of the comb.

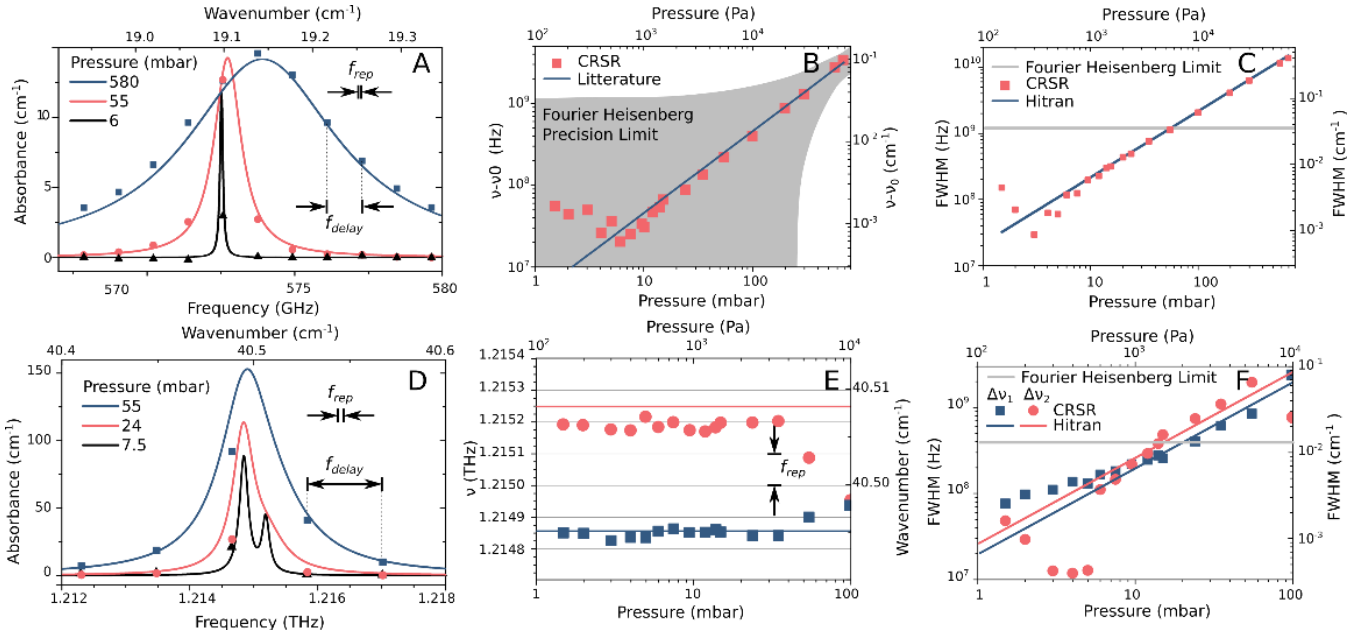

Fig. 3. CRSR analysis of NH_3 THz lines. Panel A, shows the reconstructed 573 GHz lines in solid line superimposed to the usual FFT spectrum (symbols). On panel B, the centered line frequencies retrieved by CRSR are compared to the ones determined with the line frequency at P=0 and the self-shift coefficient tabulated in HITRAN [27]. The gray area depicts precision limitation of the Fourier limitation (it is calculated by adding, for the upper bound, and subtracting, for the lower bound, the Fc=1.2 GHz to the tabulated value). The panel C shows the linewidth versus the pressure. Panel D and F are similar to panel A and C respectively for the doublet of NH_3 rotational lines around 1.21 THz (no pressure self shift has been observed for the two components of the doublet). Panel E shows the central frequencies of the two lines of the doublet versus pressure compare to the literature values.

IV. Conclusion

To summarize, we implemented a super-resolution algorithm for THz-TDS and tested it on a broadband sparse gas phase spectrum exploiting the wealth of information lying in the THz-TDS time traces. This was achieved on rotational transitions of NH_3 on a spectrum spanning more than a decade. Our method delivers accurate reconstruction despite minimal algorithm optimization and the moderate experimental integration time leading to 13 minutes experiments compared to days when using FTIR with the same resolution and number of scan [30]. The CRSR method allows to overcome the *Fc* limitation of 1.2 GHz with the ability to determine line frequency and line width with accuracies better than 20 MHz and to retrieve in a large range of pressure the individual line parameters of a doublet separated by 393 MHz. The limitations of CRSR come mainly from the experimental signal to noise ratio and are thus limited by the system used. In absolute terms, an ultimate resolution limit is represented by the linewidth of the repetition rate of the laser frequency comb (which, in our case, was around 30 Hz at 100 MHz), but practically the signal to noise ratio limits the resolution to a fraction of the repetition rate of the pulsed excitation laser.

In principle, the methods can be adapted to any spectroscopy technics in the reciprocal domain such as THz-TDS or Michelson based Fourier transform spectroscopy (e.g. FT-IR), but would be even more efficient for spectroscopic methods based on frequency combs as dual comb. This is specifically salient for gas from the THz to the IR range where a high resolution is needed.

Most importantly, our approach relies entirely on the sparse nature of the acquired spectrum, and was obtained using a commercially available spectrometer without active stabilization of the laser. Moreover, THz-TDS technics are making progress every year because of improvements of the THz photoconductive antennas [31], the laser source used, the delay lines and the system itself. Nowadays several systems allow time sampling without mechanical delay lines using dual comb spectroscopy in asynchronous optical sampling [32]. All these improvements will increase both the bandwidth of the spectrum and the signal to noise ratio. The direct consequence will be a better super resolution enhancement factor on a broader spectrum and will soon enable to reach a MHz resolution closer to the Doppler limit.

Finally, the strong rotational lines of a large variety of polar molecules lie in the THz. Therefore, THz spectroscopy find applications from fundamental sciences specifically in astronomy (astrophysics, planetology), as well as in daily subjects in atmosphere monitoring regarding pollution and global warming or gas analysis for health purposes. The rapid analysis broadband of numerous gases requires over several THz and with a subMHz resolution for reliable species discrimination and quantification. The CRSR algorithm will provide a resolution better than f_{rep} with a simple tabletop, convenient, commercial THz apparatus and we anticipate that

it will be the key to the expansion of THz-TDS of gases.

REFERENCES

[1] G. Mourou, C. V. Stancampiano, and D. Blumenthal, "Picosecond microwave pulse generation," *Applied Physics Letters*, vol. 38, no. 6, pp. 470–472, 1981. [Online]. Available: https://doi.org/10.1063/1.92407

[2] D. H. Auston and K. Cheung, "Coherent time-domain far-infrared spectroscopy," *JOSA B*, vol. 2, no. 4, pp. 606–612, 1985.

[3] R. Sprik, I. Duling III, C.-C. Chi, and D. Grischkowsky, "Far infrared spectroscopy with subpicosecond electrical pulses on transmission lines," *Applied physics letters*, vol. 51, no. 7, pp. 548–550, 1987.

[4] D. Grischkowsky, S. Keiding, M. Van Exter, and C. Fattinger, "Far-infrared time-domain spectroscopy with terahertz beams of dielectrics and semiconductors," *JOSA B*, vol. 7, no. 10, pp. 2006–2015, 1990.

[5] D. Pashnev, T. Kaplas, V. Korotyeyev, V. Janonis, A. Urbanowicz, J. Jorudas, and I. Kašalynas, "Terahertz time-domain spectroscopy of two-dimensional plasmons in algan/gan heterostructures," *Applied Physics Letters*, vol. 117, no. 5, p. 051105, 2020.

[6] T. Ichii, Y. Hazama, N. Naka, and K. Tanaka, "Study of detailed balance between excitons and free carriers in diamond using broadband terahertz time-domain spectroscopy," *Applied Physics Letters*, vol. 116, no. 23, p. 231102, 2020.

[7] H. Lin, O. J. Burton, S. Engelbrecht, K.-H. Tybussek, B. M. Fischer, and S. Hofmann, "Through-substrate terahertz time-domain reflection spectroscopy for environmental graphene conductivity mapping," *Applied Physics Letters*, vol. 116, no. 2, p. 021105, 2020.

[8] Y. Abautret, D. Coquillat, M. Zerrad, X. Buet, R. Bendoula, G. Soriano, N. Brouilly, D. Heran, B. Grezes-Besset, F. Chazallet *et al.*, "Terahertz probing of sunflower leaf multilayer organization," *Optics Express*, vol. 28, no. 23, pp. 35018–35037, 2020.

[9] M. Van Exter, C. Fattinger, and D. Grischkowsky, "Terahertz time-domain spectroscopy of water vapor," *Optics letters*, vol. 14, no. 20, pp. 1128–1130, 1989.

[10] D. Bigourd, A. Cuisset, F. Hindle, S. Matton, E. Fertein, R. Bocquet, and G. Mouret, "Detection and quantification of multiple molecular species in mainstream cigarette smoke by continuous-wave terahertz spectroscopy," *Opt. Lett.*, vol. 31, no. 15, pp. 2356–2358, Aug 2006. [Online]. Available: http://ol.osa.org/abstract.cfm?URI=ol-31-15-2356

[11] A. Cuisset, F. Hindle, G. Mouret, R. Bocquet, J. Bruckhuisen, J. Decker, A. Pienkina, C. Bray, r. Fertein, and V. Boudon, "Terahertz rotational spectroscopy of greenhouse gases using long interaction path-lengths," *Applied Sciences*, vol. 11, no. 3, 2021. [Online]. Available: https://www.mdpi.com/2076-3417/11/3/1229

[12] C. Janke, M. Först, M. Nagel, H. Kurz, and A. Bartels, "Asynchronous optical sampling for high-speed characterization of integrated resonant terahertz sensors," *Optics letters*, vol. 30, no. 11, pp. 1405–1407, 2005.

[13] T. Yasui, Y. Iyonaga, Y.-D. Hsieh, Y. Sakaguchi, F. Hindle, S. Yokoyama, T. Araki, and M. Hashimoto, "Super-resolution discrete fourier transform spectroscopy beyond time-window size limitation using precisely periodic pulsed radiation," *Optica*, vol. 2, no. 5, pp. 460–467, May 2015. [Online]. Available: http://www.osapublishing.org/optica/abstract.cfm?URI=optica-2-5-460

[14] V. A. Mandelshtam and H. S. Taylor, "Harmonic inversion of time signals and its applications," *The Journal of Chemical Physics*, vol. 107, no. 17, pp. 6756–6769, 1997. [Online]. Available: https://doi.org/10.1063/1.475324

[15] V. Mandelshtam, "Fdm: the filter diagonalization method for data processing in nmr experiments," *Progress in Nuclear Magnetic Resonance Spectroscopy*, vol. 38, no. 2, p. 159, 2001.

[16] G. Hergert, J. Vogelsang, F. Schwarz, D. Wang, H. Kollmann, P. Groß, C. Lienau, E. Runge, and P. Schaaf, "Long-lived electron emission reveals localized plasmon modes in disordered nanosponge antennas," *Light: Science & Applications*, vol. 6, no. 10, p. e17075, 2017.

[17] S. Mallat, *A wavelet tour of signal processing*. Elsevier, 1999.

[18] M. R. Wall and D. Neuhauser, "Extraction, through filter-diagonalization, of general quantum eigenvalues or classical normal mode frequencies from a small number of residues or a short-time segment of a signal. i. theory and application to a quantum-dynamics model," *The Journal of chemical physics*, vol. 102, no. 20, pp. 8011–8022, 1995.

[19] S. Kawata, K. Minami, and S. Minami, "Superresolution of fourier transform spectroscopy data by the maximum entropy method," *Applied optics*, vol. 22, no. 22, pp. 3593–3598, 1983.

[20] P. Sidorenko, E. Osherovich, Y. Shechtman, Y. C. Eldar, M. Segev, and O. Cohen, "Super-resolution spectroscopy by compact representation," in *Frontiers in Optics*. Optical Society of America, 2012, pp. FM3F–5.

[21] E. Haacke, Z. Liang, and S. Izen, "Constrained reconstruction: a superresolution, optimal signal-to-noise alternative to the fourier transform in magnetic resonance imaging." *Medical physics*, vol. 16, no. 3, pp. 388–397, 1989.

[22] R. Peretti. (2018) fit@tds github repository. [Online]. Available: https://github.com/THzbiophotonics/Fit-TDS

[23] R. Peretti, S. Mitryukovskiy, K. Froberger, A. Mebarki, S. Eliet, M. Vanwolleghem, and J.-F. Lampin, "Thz-tds time-trace analysis for the extraction of material and metamaterial

parameters," *IEEE Transactions on Terahertz Science and Technology*, vol. 9, no. 2, pp. 136–149, March 2019.

[24] J. Xu, T. Yuan, S. Mickan, and X.-C. Zhang, "Limit of spectral resolution in terahertz time-domain spectroscopy," *Chinese Physics Letters*, vol. 20, no. 8, p. 1266, 2003.

[25] P. Jansen and R. Perez, "Constrained structural design optimization via a parallel augmented lagrangian particle swarm optimization approach," *Computers & Structures*, vol. 89, no. 13, pp. 1352 – 1366, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/-S0045794911000824

[26] D. Kraft *et al.*, "A software package for sequential quadratic programming," 1988.

[27] I. E. Gordon, L. S. Rothman, C. Hill, R. V. Kochanov, Y. Tan, P. F. Bernath, M. Birk, V. Boudon, A. Campargue, K. Chance *et al.*, "The hitran2016 molecular spectroscopic database," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 203, pp. 3–69, 2017.

[28] S. Belov, A. Krupnov, V. Markov, A. Mel'nikov, V. Skvortsov, and M. Y. Tret'Yakov, "Study of microwave pressure lineshifts: Dynamic and isotopic dependences," *Journal of Molecular Spectroscopy*, vol. 101, no. 2, pp. 258–270, 1983.

[29] TOPTICA. (2020) Terascan 780 / 1550 : Topsellers for frequency-domain spectroscopy. [Online]. Available: https://www.toptica.com/fileadmin/Editors_English/-11_brochures_datasheets/01_brochures/-toptica_BR_THz_TeraScan.pdf

[30] Bruker-GmBh. (2020) The ifs 125hr ft-ir spectrometer. [Online]. Available: https://www.bruker.com/fr/products/-infrared-near-infrared-and-raman-spectroscopy/ft-ir-research-spectrometers/ifs-125hr.html

[31] R. Kohlhaas, S. Breuer, S. Nellen, L. Liebermeister, M. Schell, M. Semtsiv, W. Masselink, and B. Globisch, "Photoconductive terahertz detectors with 105 db peak dynamic range made of rhodium doped ingaas," *Applied Physics Letters*, vol. 114, no. 22, p. 221103, 2019.

[32] A. Bartels, R. Cerna, C. Kistner, A. Thoma, F. Hudert, C. Janke, and T. Dekorsy, "Ultrafast time-domain spectroscopy based on high-speed asynchronous optical sampling," *Review of Scientific Instruments*, vol. 78, no. 3, p. 035107, 2007.

Supplementary information

I. BENCHMARK : NUMERICAL VALIDATION

For benchmarking purposes, we modeled the absorption of several pressures of $^{14}NH_3$ in the THz range with an optical path length of 7 cm as in the experiments. Here, f_{rep}=1/τ_{rep} =100 MHz and f_{delay}=1/τ_{delay}=1.178GHz. Similarly, to the experiments, we focused on three key performances. First its ability to point the center of a line. Second its ability to determine its linewidth. Finally, its ability to distinguish a doublet (two close spectral lines). The intense line around 572.5 GHz of the J=0 →1, K=0 rotational transition is used to test our two first criteria and the doublet J=1 →2, K=0 & K=1 around 1.21 THz for the last one. We convoluted an experimental reference with the model of the lines using the parameters found in the HITRAN database [27] and we added a Gaussian white noise on the resulting time traces of same magnitude as in the experiments corresponding to a dynamic range of 70 dB. We then used the CRSR on these data. The results are depicted in fig S1.

There, we compare the targeted ideal line we built from HITRAN to the retrieved one. To figure out the benefit of CRSR, we added on the same plot the usual spectrum plotted in THz-TDS in large symbols (the sampling corresponds to the τ_{delay} time window) and the same data if we could record the time trace with a time window τ_{rep}.

For the case of the isolated line at 572.5 GHz, the broad absorption feature is resolved at 60 mbar by only three f_{delay} datapoints and many f_{rep} points. The linewidth here represents the resolution limit for a standard THz-TDS instrument using a τ_{delay}time window. At 6.0 mbar, the line is some ten times narrower than f_{delay} and of similar width compared to f_{rep}. With a further reduction in pressure to 0.6 mbar the linewidth is one hundred times narrower than f_{delay} and ten times narrower than f_{rep}. At these two latter pressures, the line cannot be correctly resolved using the standard approach taking the Fourier transform the temporal recording.

The CRSR algorithm accurately retrieves the line parameters in all the cases including at the lowest pressure. Both the central frequency (including the pressure shift) and the linewidth are retrieved very precisely, validating our two first assessment objectives. On panel B, the linewidths fall below the usual resolution limit for pressures below $P_{f_{delay}}$. The reconstruction stays accurate above 0.6 mbar. This corresponds to lines much narrower than the resolution limit and even narrower than f_{rep}. At pressures below 0.6 mbar, the error becomes too high and the line is not correctly reconstructed. In fact, one can derive the total absorbed energy by the line versus the pressure P:

$$SNR \propto \sigma = P \times \frac{P}{P+P_{f_{delay}}} \quad (1)$$

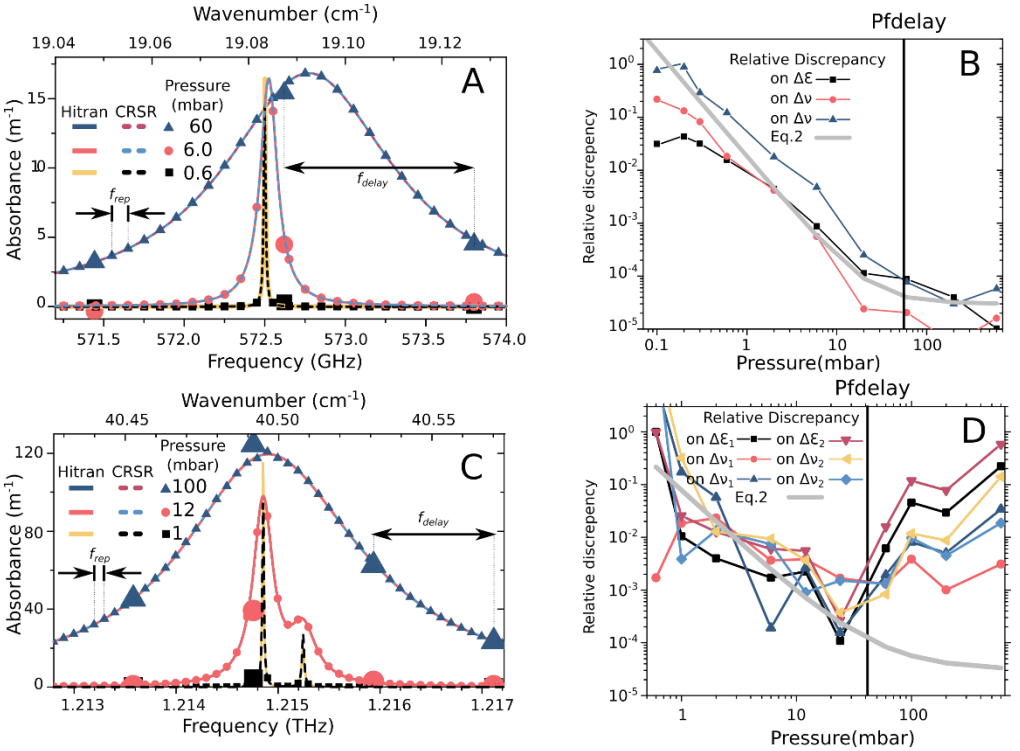

Fig S. 1. Numerical example of Super resolution: results and performances assessment. In panel A the THz-TDS spectrum of the line around 573 GHz, (large symbols) shows how the line spread spectrally using the classical mechanical delay stage. In small symbols, we plotted the FFT of a hypothetical signal where the full time (τ_{rep} time window) trace would be recorded. The reconstructed super resolution lines (dotted lines) superimposed very well with the targeted line (plain lines). In panel B, the precision of the retrieval of each parameter of the lines is plotted. It demonstrates the very good performance of the CRSR down to very low pressure. Panel C and D are similar, respectively, to panel A and B but for the doublet around 1.21 THz. Panel C demonstrates the ability of CRSR to discriminate a doublet separated by one third of f_{delay} (393 MHz). On Panel D the comparison with eq. 2 indicates that at low pressure the retrieval is limited by the signal to noise ratio when at higher pressure (when the width of the lines of the doublet are higher than the frequency separation), overlap effects deteriorate the precision of the method.

The first term comes from the number of molecules crossed by the THz beam. The second term comes from the finites size of the delay line and plays a role for lines narrower than f_{delay} (being at $P_{f_{delay}}$). In this case, the time signal extends further than the delay line windows and this fraction of the signal is lost. To analyze this more quantitatively we plotted the floored inverse function of the SNR on figure S1 panel B:

$$E_{rr} = \theta_{noise} + \frac{S_{tot}}{\sigma} \quad (2)$$

Where $S_{tot} = 3.10^{-4}$ corresponds to a fitted normalization factor to the signal (set one for all for one transition) and θ_{noise} (=3.10^{-5}, fitted as well) is the best accuracy we can reach considering the noise and the performances of the algorithm. There, the discrepancy on each parameter and the mean square error on the spectrum follows the trend of the equation only coming from the energy absorbed by the line. To conclude, we enable super resolution on lines down to about one tenth of f_{rep} and one hundredth of the usual resolution limitation for this level of noise for the two first criteria. We establish that the resolution performance of the CRSR technique is dependent on the SNR. Meaning the SNR of the molecular signal not of the source. The natural conclusion is therefore for a stronger absorber, or a longer interaction length a resolution improvement is foreseen.

The most convincing proof of super resolution is a technique capacity to resolve doublet. We examined the doublet at 1.21 THz, which exhibits a frequency difference of ~ 393 MHz and stands above f_{rep} and below f_{delay} making it a very good example. It is important to note in performing the usual FFT both lines of the doublet stand in the same frequency point as seen on panel C of fig S1. As in the single line case, we plotted the retrieved spectrum and the error on panel D. Once again, CRSR retrieves very accurately the parameters from 3 mbar to 100 mbar. However, the pressure range of reliability is narrower. At pressures below 1 mbar the doublet is not retrieved properly. More quantitatively, with equation 2 added to panel D with the same θ_{noise} and a $S_{tot} = 1/629$; to take into account the lower signal at this frequency (~8 dB compared with 572 GHz). As expected, equation 2 explains the low-pressure behavior but does not take into account the effect the overlapping lines occurring above 100 mbar. Here the lifetime of the oscillator (inverse of the linewidth) became shorter than the beating period of the doublet. Thus, the overlap degrades the accuracy of the fit. Still, we enable super resolution of a doublet down to one third of the usual resolution limitation for this level of noise. **We want to emphasize here that contrary to what we did for the experimental data, we did not need here the additional constraint of fixing the ratio between the two lines intensity**. This shows that the doublet separation is fully achievable with CRSR even without the need of additional constraint as soon as the signal to noise ratio is good enough.

Using fictitious samples of a single line and a doublet, at pressures from 0.1 mbar to atmospheric pressure, including an experimental level of white Gaussian noise, we showed the capacity of CRSR to retrieve the spectroscopic parameters of rotational lines from THz-TDS time traces. With decreasing pressure the signal is reduced, nevertheless CRSR provides accurate parameters down to 0.6 mbar for a single line corresponding to a line narrower than 30 MHz. the resolution demonstrated by this approach is easily sufficient to fully resolve the lines separated by 400 MHz at 1 mbar. This is coherent with the resolution limit observed for a single line of approximately 30 MHz. This means a super-resolution of a 30-fold factor improvement.

II. Full spectrum reconstruction at 9.5 mbar

In the main text, we focused on individual peaks/doublet to analyze the precision and the performances of the CRSR. Here, to show that our approach is broadband, we reconstructed the whole spectrum from fig 2 of the main text. For this reconstruction, we begin with the line at 573 GHz and then added the lines at higher frequency, few by few up to 4 THz (the last observable line is around 3.5 THz) without using data from Hitran except for the 0.38 factor of the doublet we already explained. The result of the reconstruction is shown on table S1. It is important to state here that the only input from Hitran remains the 0.38 factor for the 1.21 THz doublet. We did not use any other information than this factor in our whole study and thus to retrieve the following table.

Upper State			Lower State			literature values		Retrieved value		difference		Relative difference	
J'	K'	a/s	J"	K"	a/s	Freq.(THz) P=0 mbar	FWHM (MHz) P=10 mbar	Freq. (THz)	FWHM (MHz)	Freq. (MHz)	FWHM (MHz)	Freq.	FWHM
1	0	s	1	0	a	0.5725	215.395	0.57253	199.927	31.6	15.5	$5.51\ 10^{-5}$	0.0718
2	1	s	1	1	a	1.1685	258.592	1.1685	303.315	36.4	44.7	$3.11\ 10^{-5}$	0.173
2	0	a	1	0	s	1.2149	195.867	1.2148	219.763	51.3	23.9	$4.23\ 10^{-5}$	0.122
2	1	a	1	1	s	1.2152	258.592	1.2152	218.024	28.8	40.6	$2.37\ 10^{-5}$	0.157
3	**0**	**s**	**2**	**0**	**a**	**1.7635**	**178.706**	**1.7621**	**1365.58**	**1410**	**1190**	$\mathbf{8.01\ 10^{-4}}$	**6.64**
3	1	s	2	1	a	1.7636	234.330	1.7635	113.079	109	121	$6.15\ 10^{-5}$	0.517
3	2	s	2	2	a	1.7638	293.505	1.7642	5.23188	373	288	$2.11\ 10^{-4}$	0.982
3	1	a	2	1	s	1.8089	234.330	1.809	248.148	57.7	13.8	$3.19\ 10^{-5}$	0.059
3	2	a	2	2	s	1.8104	293.505	1.8104	245.363	30.2	48.1	$1.67\ 10^{-5}$	0.164
4	3	s	3	1	a	2.3572	213.027	2.3574	296.553	166	83.5	$7.04\ 10^{-5}$	0.392
4	2	s	3	2	a	2.3577	265.101	2.3583	243.703	593	21.4	2.52E-4	0.0807
4	3	s	3	3	a	2.3586	320.133	2.3588	14.1798	237	306	1E-4	0.956
4	**0**	**a**	**3**	**0**	**s**	**2.4000**	**163.913**	**2.4**	**3**	**3.4**	**161**	$\mathbf{1.42\ 10^{-6}}$	**0.982**
4	1	a	3	1	s	2.4006	213.027	2.4004	565.219	154	352	$6.4\ 10^{-5}$	1.65
4	2	a	3	2	s	2.4023	265.101	2.4023	155.681	29.2	109	$1.21\ 10^{-5}$	0.413
4	3	a	3	3	s	2.4051	320.133	2.4051	312.653	0.42	7.48	$1.75\ 10^{-7}$	0.0234
5	0	s	4	0	a	2.9484	150.894	2.9486	463.737	162	313	$5.5\ 10^{-5}$	2.07
5	1	s	4	1	a	2.9487	193.500	2.949	543.771	261	350	$8.84\ 10^{-5}$	1.81
5	2	s	4	2	a	2.9495	239.064	2.9493	451.308	233	212	$7.89\ 10^{-5}$	0.888
5	3	s	4	3	a	2.9508	287.587	2.9508	420.287	0.28	133	$9.49\ 10^{-8}$	0.461
5	4	s	4	4	a	2.9526	339.069	2.9527	33.2146	106	306	$3.59\ 10^{-5}$	0.902
5	1	a	4	1	s	2.9896	193.500	2.9897	288.323	53.9	94.8	$1.8\ 10^{-5}$	0.49
5	2	a	4	2	s	2.9916	239.065	2.9916	167.763	7.19	71.3	$2.4\ 10^{-6}$	0.298
5	3	a	4	3	s	2.9948	287.587	2.9948	237.209	36.1	50.4	$1.2\ 10^{-5}$	0.175
5	4	a	4	4	s	2.9994	339.069	2.9995	309.077	108	30	$3.59\ 10^{-5}$	0.0885
6	1	s	5	1	a	3.5374	176.931	3.5373	47.23	92.8	130	$2.62\ 10^{-5}$	0.733
6	2	s	5	2	a	3.5385	215.394	3.5386	60.1745	86.7	155	$2.45\ 10^{-5}$	0.721
6	3	s	5	3	a	3.5403	257.408	3.5404	334.802	87.2	77.4	$2.46\ 10^{-5}$	0.301
6	4	s	5	4	a	3.5428	301790	3.5429	156.65	81.7	302000	$2.31\ 10^{-5}$	0.999
6	5	s	5	5	a	3.5460	349.721	3.546	455.652	31.2	106	$8.81\ 10^{-6}$	0.303
6	0	a	5	0	s	3.5749	140.835	3.5752	353.929	262	213	$7.32\ 10^{-5}$	1.51
6	1	a	5	1	s	3.5756	176.932	3.5754	645.983	170	469	$4.74\ 10^{-5}$	2.65
6	2	a	5	2	s	3.5777	215.395	3.5778	284.482	111	69.1	$3.09\ 10^{-5}$	0.321
6	3	a	5	3	s	3.5813	257.409	3.5813	413.398	14.3	156	$4\ 10^{-6}$	0.606
6	4	a	5	4	s	3.5865	301.790	3.5865	245.971	24.8	55.8	$6.91\ 10^{-6}$	0.185
6	5	a	5	5	s	3.5933	349.721	3.5933	700.185	13.6	350	$3.78\ 10^{-6}$	1

Table S 1 : Value of the retrieved parameter for the whole spectrum at 10 mbar. The multiplets (lines closer than the usual resolution of 1.2GHz) are shown by using a gray background. In red, we show the lines where the retrieval is worse than the resolution or limited by the methods (reaching the lower bound for the linewidth). FWHM: Full-Width at Half Maximum

Here one can see that we retrieved the line central frequency and the width better than the resolution for all the lines except one (35 over 36). The second criteria was met at a precision of 50 % for two third of the lines (24 over 36). It shows that the methods stay valid on most of the spectrum. Still a lack of accuracy exist when dealing with a complex multiplet because more information is needed for the same signal to noise ratio as the one around 1.76 THz. Additionally, when going at high frequency the signal to noise ratio decrease, the precision decrease as well. It shows that the methods is valid on the whole spectrum and clearly benefit from a high signal to noise ratio. Fig S3 illustrates the

sensitivity to the J- and K-dependence.

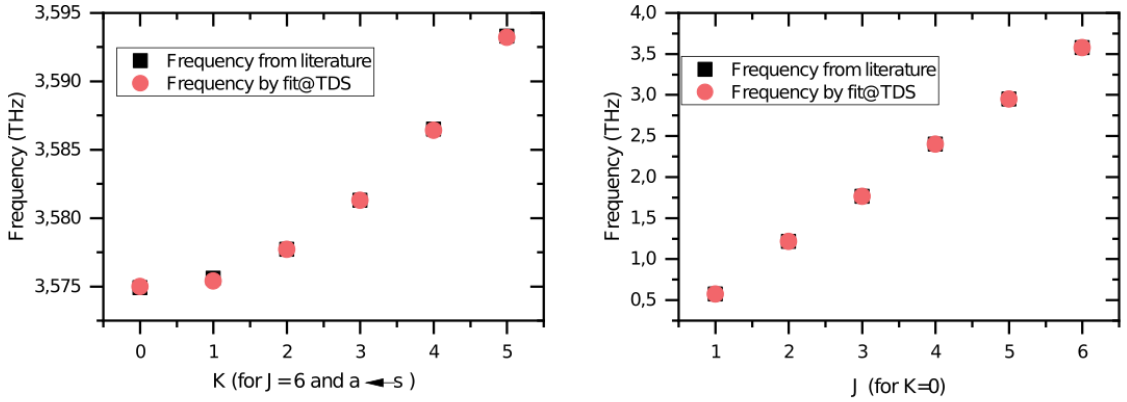

Figure S2: K-dependence (left) and J-dependence (right) for the frequency retrieved by the CRSR method

Overall, it shows that the approach is fully broadband and only limited by noise of the spectroscopic set up to the signal analysis.

III. Group of 4 lines around 3 THz

In the main text, we showed the performances of the CRSR on lines where the THz-TDS spectrum has most energy. To show that our approach is applicable on the whole spectrum and that the performances will only depend on the signal to noise ratio, we show how the CRSR can retrieve the parameters from the lines around 3 THz. The results are plotted on fig S4 comparable to fig 3 of the main text.

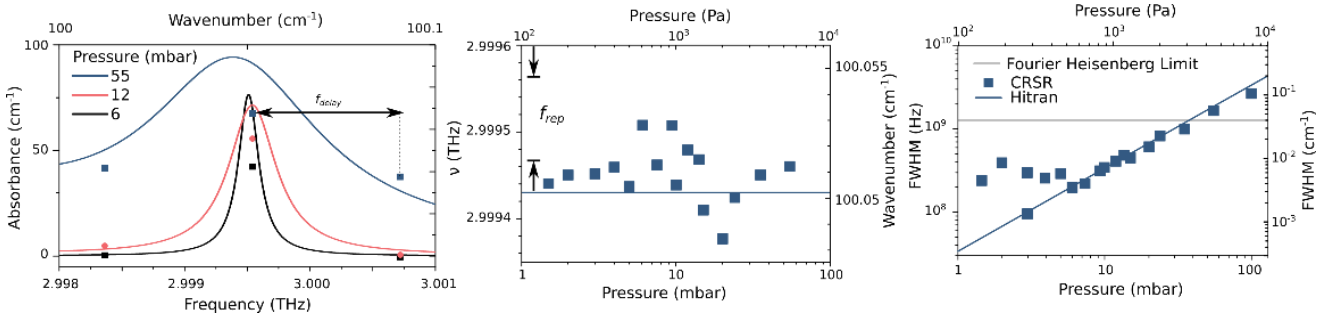

Figure S3: Super resolution for NH_3: results and performances assessment, additional data around 3 THz. Panel A, shows the reconstructed 2.999 THz lines in plain line superimposed to the usual FFT spectrum demonstrating the super resolution reconstruction. Panel B illustrates that the frequency of the line is precisely retrieved (±100 MHz) for pressure from 2 mbar to 100 mbar. Panel C shows the linewidth of the line versus the pressure demonstrating precise retrieval for pressure from 6 mbar to 100 mbar.

Here one can see that we are able to reach performance comparable to the one obtained at higher energy and conclude that the methods is valid on the broad THz-TDS spectrum.

IV. Absolute frequency normalization on water lines

In a THz-TDS system, the frequency quantification and its precision comes from quantification and the precision of the time sampling step that is by itself tainted with error [2]. The manufacturer corrected these measurements for systematics error to reduce it to the minimum and to get an as good as possible absolute value of the time delay. However, this cannot be done at the precision needed for super resolution, simply because CRSR is not implemented yet on the commercial system . To compensate for this, we recorded the spectrum of water vapor around 570 GHz, and we used the CRSR to get the central frequency of this very known line. We used the ratio found (1.00086423) between the database value and the retrieved value to normalize the value of our time step and get better absolute measurement. We estimated the precision on the time step to be better than 4 10^{-5} from the precision on the NH_3 lines central frequency.

V. Stability of the repetition frequency of the fiber laser

We measured the stability of repetition rate of the laser to be sure this will not be a limitation for our experiments. We found a long time deviation of 30 Hz on the repetition rate that is relative to a precision of 3 10^{-7}, or a variation in repetition time of 3 fs. This is in good agreement with the 1.405 fs of the delay noise of two consecutive experiments. It is important to note that this jitter is below 10 % of the time step (100/3 of fs). It means that for narrow line with time spanning wider than the repetition time, the measurement will still be in the good time step. Said differently, one can sample the signal originating from one pulse of the laser with the next pulse and still be in the good time step. It is an important prerequisite if we want to

reach super resolution better than the repetition frequency.

VI. Other supportive information

The reader will find the source code for the super resolution, the one to create the numerical examples, the numerical examples we created as the data from the NH_3 experiment in zip files linked to this paper.

.

VII. References

[1] Iouli E Gordon, Laurence S Rothman, Christian Hill, Roman V Kochanov, Y Tan, Peter F Bernath, Manfred Birk, V Boudon, Alain Campargue, KV Chance, et al. The hitran2016 molecular spectroscopic database. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 203:3–69, 2017.

[2] Arno Rehn, David Jahn, Jan. C. Balzer, and Martin Koch. Periodic sampling errors in terahertz time-domain measurements. *Opt. Express*, 25(6):6712–6724, Mar 2017.