# Supplementary Information to "Computational Doppler-limited dual-comb spectroscopy with a free-running all-fiber laser"

**Łukasz A. Sterczewski[1], Aleksandra Przewłoka[2], Wawrzyniec Kaszub[2], Jarosław Sotor[1,*]**
[1]*Laser & Fiber Electronics Group, Faculty of Electronics, Wroclaw University of Science and Technology, Wybrzeze Wyspianskiego 27, 50-370 Wroclaw, Poland*
[2]*Institute of Electronic Materials Technology, Wolczynska 133, 01-919 Warsaw, Poland*
*Author e-mail address: jaroslaw.sotor@pwr.edu.pl*

## Detailed Methods:

**Detailed dual-comb generation setup.** In the dual-comb generation setup, we used a fiber-coupled pump diode with a center wavelength of 976 nm supplied with 250 mA of current and a voltage bias of 1.43 V (SK17511049) provided by a laser driver (CLD1015, Thorlabs). Under such bias conditions, the diode emitted ~119 mW of optical power corresponding to ~33% of wall-plug efficiency, and this parameter predominantly leaves room for improvement of the power budget in future implementations. The pump light was coupled to the ring oscillator through a single-stage Tap/WDM/Isolator device (TIWDM-980/1550S-10%-250, Opneti) introducing 8.8% of losses from the pump input port to the output. The TIWDM served also as a 10% output coupler. It was spliced to a 26 cm long piece of low-dispersion erbium doped fiber (EDF150LD, OFS Optics) with a peak absorption of 150 dB/m and a dispersion of −16 ps/(nm·km), further followed by a graphene-based saturable absorber (SA). To fabricate the absorber, 20 layers of graphene were deposited on a poly(methyl methacrylate) (PMMA) substrate using procedures described in Ref.[35], and next the substrate with the graphene was transferred onto the facet of an angle-polished fiber connector. Finally, the absorber was sandwiched between a complementary connector so that its center was located ~72 cm from the output tap (measured to the center of the hybrid component). The cavity included also an in-line polarization controller (F-POL-IL, Newport) preceding an 18 cm long piece of panda-type polarization maintaining fiber optimized for 980 nm (PM980, Nufern) located 28 cm away from the hybrid component. The total cavity length was ~1.40 m, and except for the PM section, it utilized an ordinary 1550 nm single mode fiber (SMF-28e, Corning). The use of the 980 nm PM section rather than optimized for 1550 nm was dictated by a nearly 50% decreased beat length, which enabled stable dual-comb operation despite the short cavity length allowing to reach higher repetition rates more suited for molecular spectroscopy. We estimate 3.5 dB of loss in the fiber laser cavity based on the insertion losses of the TIWDM and the saturable absorber. The cross-polarized dual-comb light was collected directly from the tap of the hybrid device without any additional optical or optomechanical components. The mode-locked operation of the laser was only possible when the graphene/PMMA composite was placed onto the fiber connector. In the absence of the composite, we did not observe any sign of mode synchronization through the weak artificial saturable absorber realized by the piece of PM fiber and polarization controller. The polarization multiplexing mechanism is in fact strongly supported by the polarization insensitive graphene-based saturable absorber.

**Detailed dual-comb spectroscopy setup.** The output of the dual-comb oscillator was guided to a cascade of two fiber-coupled, 16.5 cm long absorption cells containing hydrogen cyanide ($H^{13}C^{14}N$) at a low pressure of 10 Torr. Next, it was followed by an in-line polarization controller (Fibrepro, PC1100) and a fiber-based polarizing beam splitter (PBS, 1 SM, 2 PM ports, >20 dB extinction ratio). One of the fiber outputs was connected to a fast extended InGaAs photodetector (DCS-50S-39, Discovery Semiconductors) to produce a radio-frequency (RF) dual-comb interferogram. Its electrical output was split into two paths with a resistive RF power splitter for simultaneous measurement of the dual-comb RF intermode beat notes with an RF spectrum analyzer (EXAN9010A, Agilent) and acquisition of dual-comb interferograms, yet in an actual device that splitting may not necessary and served only for diagnostic purposes, just like a measurement of the optical spectrum using an optical spectrum analyzer (AQ6375, Yokogawa, 50 pm resolution) from the other port of the PBS. It should be underlined, however, that in the polarization-resolved characterization (Fig. 2 of the main manuscript), the optical spectrum RF spectra were measured from the same PBS output after splitting with a 50/50 PM fiber coupler.

The electrical output of the photodetector that was used for measuring dual-comb interferograms required electrical amplification and filtering to avoid RF spectral aliasing. We used a home-made 70 MHz LC-type low-pass filter followed by a 40 dB low-noise (0.8 dB noise figure) RF amplifier (PE15A1012, Pasternack), and another low-pass filter prior to sampling the electrical signal with a fast 8-bit oscilloscope (Infiniium DSO90604A, Agilent). We used the oscilloscope in oversampling (high vertical resolution) mode by taking advantage of its high sampling rate (20 GS/s) to acquire digital interferograms at 400 MS/s with 11 effective number of bits (ENOB).

**CW laser experiment.** To characterize the relative and absolute linewidths of the dual-comb laser, we optically heterodyned a selected portion of the dual-comb spectrum with a narrow-linewidth continuous wave (CW) laser (Ando AQ4320D, ~200 kHz linewidth). The fiber-coupled source operating at a center wavelength of ~1560 nm and an optical power of 1 mW was combined with comb lines lying in the central part of the optical spectrum filtered using a home-made tunable optical band-pass filter. We used a non-PM 90/10 optical coupler to guide most of the low-intensity dual-comb light to the 90% coupling ratio port, whereas the 10% port was used for the CW laser acting as a strong local oscillator. The optical beating signal was measured using the previously mentioned extended InGaAs photodetector connected to the oscilloscope (together with the instantaneous repetition rate signal and dual-comb interferograms from a separate detector unit). The CW laser mixing signal was filtered and amplified using the same RF circuitry as in the Detailed Spectroscopy Setup. After acquisition, the measured signal was subject to line width analysis.

First, we generated a waterfall plot (RF spectrogram in Fig. 3a) of the beating signal between two comb lines, one from each comb, and the CW laser, which was calculated with a temporal resolution of 20 µs and an intensity threshold of -55 dB. It reveals highly correlated fluctuations of the two comb lines. From a dual-comb spectroscopy perspective, of particular interest is the relative comb linewidth, hence to characterize it, we digitally squared the CW laser mixing signal to induce self-mixing between the two isolated comb lines, thus giving rise to a strong RF line located exactly at the difference frequency between them. The line was so narrow that for the generation of the difference-frequency spectrogram in Fig. 3b, we picked a frequency resolution of 3 kHz, and 70% overlap between the periodogram blocks at the expense of a relatively sparse temporal grid. To ultimately prove the high coherence between the two RF lines, we calculated the Fourier Transform on the full-length self-mixing signal, which yielded the 200 Hz relative line width (at 100 Hz of resolution bandwidth) quoted in text.

The estimate of the absolute line width of the comb lines was obtained in two ways. First, we calculated the persistence spectrum of the lower-frequency beat note centered at ~12 MHz in a bandwidth of ~14 MHz with a frequency resolution of 1.56 MHz, and a temporal resolution of 20 µs, as plotted in the top panel Fig. 3d. This type of plot is well-suited for analysis of non-stationary signals, i.e. such that change frequency characteristics over time. Sporadic excursions over a wider frequency range, as well as the bandwidth of long-time persistence are clearly visible in the figure. For fair comparison, however, we perform a conventional Fourier-based analysis on the entire acquired signal, as plotted in the bottom panel of Fig. 3d. It should be noted, however, that the estimated linewidth is in fact a convolution of the profiles of the comb lines and the CW laser, hence the quoted number is inevitably overestimated. Nonetheless, such megahertz-scale fluctuations can be neglected when compared to Doppler linewidths of typical molecular transitions at room temperature.

**All-computational phase correction algorithm.** Dual-comb interferograms are predominantly corrupted by two kinds of noise: additive amplitude noise and multiplicative phase noise. Provided that one can deal with the latter, amplitude fluctuations responsible for uncertainty of the spectral transmittance can be minimized through prolonged averaging. Unfortunately, phase noise cannot be suppressed using trivial techniques, requiring sophisticated carrier and timing recovery algorithms instead. While such are widely developed for wireless network modulation schemes, the field of computational multiheterodyne spectroscopy[37] is just emerging. And it is not the correction that is difficult *per se*, but the retrieval of phase and frequency fluctuations. Unlike techniques relying on external CW lasers to extract the correction signals, our solution is all-computational and operates on raw time-domain interferograms. Rather than relying on extremely effective yet computationally demanding cross-correlation or cross-ambiguity functions[38,44] or the Extended Kalman Filter (EKF)[37,45] to simultaneously extract the RF carrier frequency $f_c$ and RF pulse repetition rate $\Delta f_{rep}$, we took a radically different approach. We decomposed the problem into two sub-problems: carrier-frequency-independent estimation of the interferogram's time-of-arrival followed by frame-to-frame carrier phase tracking.

We take advantage of the characteristic temporal structure of the dual-comb signal. Because there is a strong centerburst in the interferogram, we can use it as a trigger to measure the quasi-instantaneous RF repetition rate $f_{rep}$. Unfortunately, ordinary threshold (leading edge) triggering introduces a considerable timing jitter and shows high sensitivity to carrier phase. To make the triggering scheme virtually insensitive to carrier parameters, we can calculate the envelope of the signal either by quadrature demodulation of the RF dual-comb signal or by applying the Hilbert Transform to obtain an analytic representation. In a real-time platform, a finite-impulse-response (FIR) approximated Hilbert transform filter can be implemented, which is now supported by most hardware architectures. In principle, this operation is similar to amplitude demodulation of a signal. Mathematically, a real (passband) dual-comb signal $y(t)$ with an instantaneous carrier frequency $f_c(t)$ can be converted into the analytical representation through

$$y_a(t) = y(t) + \mathrm{i} \cdot \mathcal{H}(y)(t) = A(t) \cdot \exp\left(\mathrm{i}\left(\varphi_\mathrm{c}(0) + 2\pi \int_0^\mathrm{t} f_c(\tau)\mathrm{d}\tau\right)\right), \quad (1)$$

where $A(t)$ is the instantaneous amplitude of the envelope, $\mathcal{H}$ is the Hilbert transform, and $\varphi_\mathrm{c}(0)$ is the initial phase. By taking the absolute value of $y_a(t)$, we obtain a carrier-frequency-independent signal, which after low-order low-pass (FIR) filtering yields a good estimate of the true instantaneous amplitude. Alternatively, one can use hardware-based quadrature (IQ) demodulation to obtain the same. That signal is next fed to a constant fraction discriminator (CFD) block, which in contrast to leading-edge filtering does not introduce timing jitter and is more robust than ordinary maximum peak trigger failing to be accurate for signals without a sharp maximum. This technique is widely used in scintillation counters and relies on triggering at a level being a fraction of the peak height with a theoretical zero time walk. The CFD basically consists of an ordinary delay, negation, summation and zero crossing detector block (see Supplementary Note 7). The input signal is divided into two, one of which is inverted, time delayed by $\tau_\mathrm{D}$, and added to the other. The resulting waveform has a well-defined zero crossing suitable for triggering. In other words, we find such $t$ around the centerburst where the sign of $A(t) - A(t - \tau_\mathrm{D})$ changes from positive to negative. Of course, the *a priori* known delays introduced by FIR filtering and the delay block of the CFD should be compensated for accurate determination of the time of arrival. Next, we divide the signal into individual interferogram frames, and derive the discrete (frame-by-frame) repetition rate $\Delta f_\mathrm{rep}[i]=1/\Delta T_\mathrm{rep}[i]$, followed by adaptive resampling[43,39] using a cubic polynomial to ensure the equal duration of the frames. In the repetition correction we assume that the repetition rate changes linearly from the center of the interferogram on the left $\Delta f_\mathrm{rep}[i-1]$ to the center of the corrected frame with a repetition rate $\Delta f_\mathrm{rep}[i]$, and next changes linearly to the center of the interferogram on the right with $\Delta f_\mathrm{rep}[i+1]$, which yields a continuously varying $\Delta f_\mathrm{rep}(\mathrm{t})$. Next, we provide an additional phase shift[52] induced by the timing jitter $T_\mathrm{jitter}(t)$, which is obtained by multiplying the complex resampled interferogram frame $y_\mathrm{r}(t')$ by $\exp(\mathrm{i}2\pi f_\mathrm{c} T_\mathrm{jitter}(t'))$, thus yielding $y_\mathrm{rc}(t')$. Now all the comb teeth share common frequency fluctuations, as visible in the spectrum of the digital difference frequency (DDFG) signal[39] (see Supplementary Note 7, Supp. Fig. 8), and thus any dependence of phase noise characteristics on the mode order number is virtually eliminated. The only remaining fluctuation is common phase noise. Rather than calculating the absolute phase around the center of each interferogram, we calculate the phase increment between the corresponding samples around the center of the current ($i$) and previous ($i-1$) interferogram frame. For the discrete time case that is $y_\mathrm{int}[i,n] = y_\mathrm{rc}[i,n]y_\mathrm{rc}^*[i-1,n]$, where the asterisk denotes complex conjugate, and $n$ is the sample number. Next, we obtain the mean of the complex samples $y_\mathrm{int}[i,n]$, and retrieve the (mean) discrete-time advancement of phase between two consecutive interferogram frames $\Delta\varphi_\mathrm{c}[i] = \arg\{\bar{y}_\mathrm{int}[i]\}$. Finally, a linear phase correction vector $\varphi_\mathrm{cv}$ is constructed by the same way as in the repetition rate case (phases in the center of the neighboring left $\Delta\varphi_\mathrm{c}[i-1]$, current $\Delta\varphi_\mathrm{c}[i]$, and neighboring right $\Delta\varphi_\mathrm{c}[i+1]$ interferograms are known, and they are assumed to vary in a linear fashion). Phase correction of the of $i$-th interferogram frame therefore given by

$$y_\mathrm{cor}[i,n] = y_\mathrm{rc}[i,n] \cdot \exp\left(-\mathrm{i}\left(\varphi_\mathrm{cv}[i,n] + \left(\sum_{k=1}^{i-1} \Delta\varphi_\mathrm{c}[k]\right) \bmod 2\pi\right)\right). \quad (2)$$

Note that this phase correction ensures equal phases of all interferogram frames, hence it turns the RF comb into a harmonic one, i.e. without a carrier-envelope offset. Despite the apparent complexity, the algorithm operates in an all frame-by-frame architecture with no significant computational requirements. Repetition rate correction is basically a linear (or cubic) interpolation problem, whereas phase carrier extraction and correction are nothing but a calculation of the phase angle and complex multiplication supported by the CORDIC algorithm. Consequently, the extraction of the correction parameters is fast and bodes for future real-time implementations.

The proposed digital phase correction method, however, has several limitations, of which the most critical is the necessity of providing sufficient mutual coherence between the combs on a $1/\Delta f_\mathrm{rep}$ scale. In other words, large and regular carrier phase jumps beyond $\pi$ may lead to incorrect spectroscopic results. Also, an accumulation of numerical errors may lead to a drifting carrier frequency and phase. It is possible to address this potential issue by introducing a second phase correction step comparing the current frame with the first acquired interferogram or the average of those already accumulated. Another challenge is the discrete nature of acquired samples implying roundoff errors in the zero-crossing detection for repetition rate estimation because the result must be an integer number. Higher sampling rates or digital interpolation filters will definitely improve this procedure. Finally, the proposed method improves only the linewidth of RF comb lines, while leaving optical domain fluctuations intact.

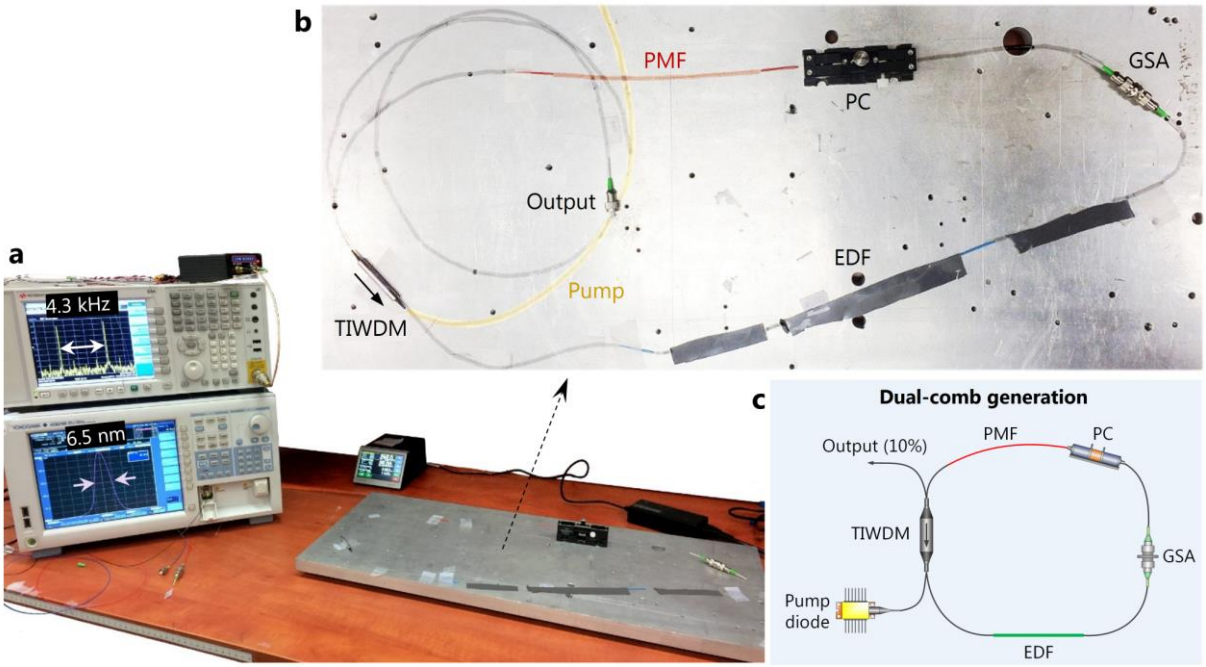
**Spectroscopy of HCN.** The measured isotope of hydrogen cyanide $H^{13}C^{14}N$ has proven its usefulness for precise wavelength calibration of optical instruments in the near-infrared, yet it is still lacking a comprehensive spectroscopic line-by-line model, as opposed to the most-abundant isotope $H^{12}C^{14}N$ conveniently available in the HITRAN[53] database. To demonstrate the precision and accuracy of our spectrometer, we compared the dual-comb measurement fit with a tunable diode laser spectroscopic scan (TDLAS), and the two spectroscopic models of the available $P$-branch transitions as described in the text. We simulated the transmission spectra using formulae in

Ref.[53] for the three available lines modeled using a Voigt profile at a temperature of 296 K with theoretical line intensities from the ExoMol database[48], and those obtained experimentally[30], whereas the pressure broadening coefficients common for both models were obtained from Ref.[47]. To account for the isotopic ratio of HCN in the cell, previously reported line intensities were scaled by the natural isotopic abundance from HITRAN[53]. In the simulation, however, we had to account for the total uncertainty budget from numerous contributors of the model, which yielded shaded-area curves representing $2\sigma$ (95%) confidence intervals. Here, the quoted uncertainty is the extended uncertainty with a coverage factor of $k$=2 (uncertainty is $\pm 2\sigma$) induced by the dominant sources of uncertainty: pressure of gas in the cell, pressure broadening coefficients, and line intensities, whereas the molar fraction, isotopic purity, cell length, and temperature were assumed to be known exactly for simplicity.

In the Voigt profile fitting routine, we used the nonlinear trust region reflective algorithm and limited the number of points being fit to the central part of 20 points around the peak: 10 on each side, which corresponds to approximately 7 Doppler widths. This was dictated by the presence of weak adjacent lines from hot-band transitions, which would otherwise corrupt the fit[47]. The Doppler widths were known a priori and fixed, while the line intensity, Lorentzian width, and center position were free parameters for the fitting routine for each line. To obtain the uncertainties of the Lorentzian line widths from the fits, we calculated them from the Jacobian estimate and the variance of residuals.

## Supplementary Note 1: Photography of the experimental setup

Supplementary Figure 1 shows a photography of the dual-comb laser used in the experiment together with the previously shown dual-comb generation block diagram for comparison. Because of the optical fibers' transparency, they were manually highlighted in panel b) to improve their visibility.

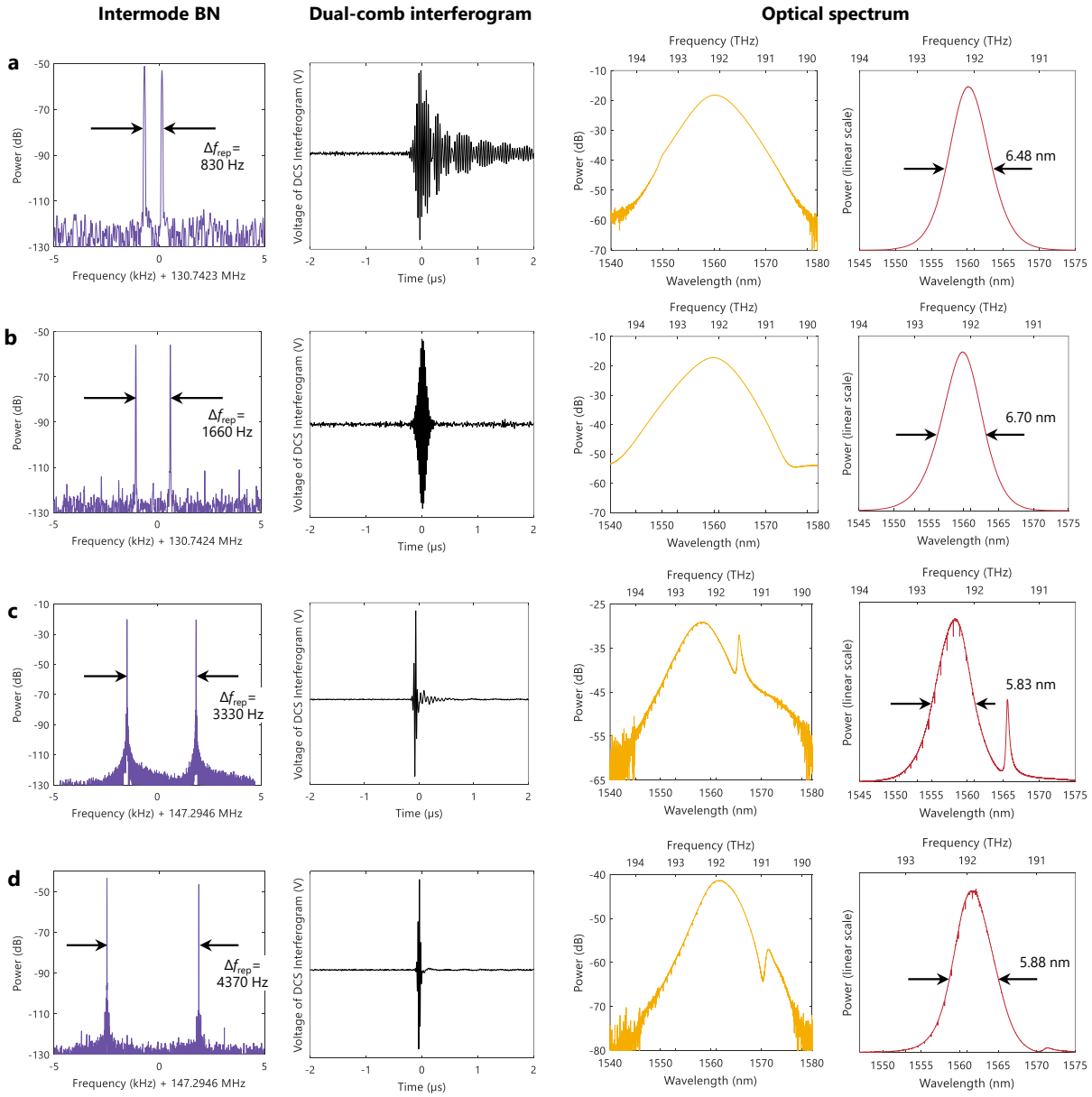



**Supplementary Fig. 1** Photography of the experimental setup along with the previously shown schematic diagram. (a) Testbed for the dual-comb laser. An optical and radio frequency analyzer (via photodetector) are connected to the output of the TIWDM. (b) Zoom of the dual-comb oscillator with manually highlighted fibers. (c) Schematic of the dual-comb generator for comparison. TIWDM – polarization insensitive hybrid Tap/Isolator/WDM component, EDF – erbium doped fiber, GSA – graphene saturable absorber, PC – polarization controller, PMF – polarization maintaining fiber.

## Supplementary Note 2: Repetition rate difference tunability

The single-cavity dual-comb laser showed a high degree of flexibility regarding the obtainable repetition rate difference, which was critical to permit accurate computational correction of Doppler-limited absorption lines. In general, the range of repetition rate difference $\Delta f_{\mathrm{rep}}$ is determined predominantly by the length of the PM fiber

section, which enables to engineer the laser for a specific application, i.e. for broadband alias-free analysis with lower $\Delta f_{rep}$, and for high-refresh-rate combustion diagnostics with $\Delta f_{rep}$ closer to 10 kHz. At a constant length of the PM section, the repetition rate could be varied using the polarization controller by as much as a factor of 2. Supplementary Fig. 2 shows two configurations of the laser with the length of the PM section equal to 14 cm, and 18 cm, and the mean repetition rates of 130.74 MHz, and 147.29 MHz, respectively. In the figure, we intentionally included dual-comb interferograms because we observed operation regimes with two beat notes separated by as low as ~200 Hz which did not produce any spectroscopic interference. Instead of two coexisting combs, the behavior of the laser resembled a rapid switching between two orthogonal polarizations which produced a continuous-wave-like signal in the radio frequency.



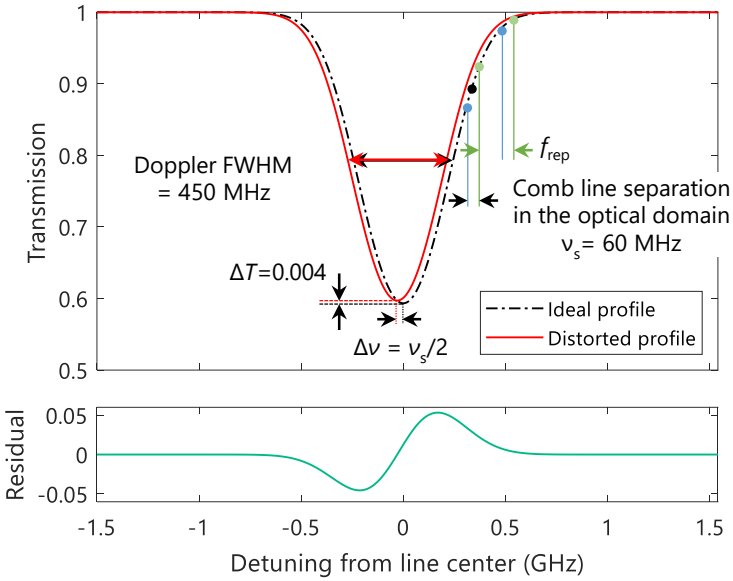

**Supplementary Fig. 2** Tunability of the repetition rate difference using the polarization controller and by varying the length of the PM fiber section. The first column plots the dual-comb intermode beat notes from the RF spectrum analyzer, second shows the dual-comb interferogram recorded with the oscilloscope, whereas the third and fourth show the optical spectrum in a logarithmic and linear scale. (a,b) Data for the 130 MHz laser. (c,d) Data for the 147 MHz laser.

It should be also mentioned that repetition rate differences exceeding 10 kHz were possible with a PM fiber longer than 30 cm (not shown), albeit with a fundamental repetition rate closer to 100 MHz.

## Supplementary Note 3: Propagation of two combs through an absorption cell

The collinear configuration of the dual-comb spectrometer implies that the optical spectrum is simultaneously probed by two frequency combs with teeth at different optical frequencies. Rather than accessing the intensity around a single comb tooth, we will measure the geometric mean between the teeth intensities of the local oscillator and the signal beam. While it is possible to split the polarizations and guide only one of the combs through an absorption cell, from a practical standpoint spatial overlap between reduces the complexity of the spectroscopic system.

To estimate the influence of this effect on the measured spectrum, we simulated a Doppler-limited transition of HCN at room temperature around 1560 nm with a 450 MHz-wide Gaussian profile (FWHM) and a peak absorption (1-*T*) of 0.4, as plotted in Supplementary Fig. 3. In the dual-comb spectrum, the optical frequency difference between the teeth of the signal and local oscillator comb increases linearly, and in the most extreme case reaches values close to a half of the fundamental repetition rate $f_{rep}$. We found that for the highest-frequency lines observed in our spectrum ($\nu_S$~60 MHz), the collinear (symmetric) configuration causes the lines to be shifted in peak position by $\nu_S/2$ (30 MHz) and leads to a decrease in peak absorption by $4\cdot10^{-3}$, which is less than the obtainable uncertainty in our system within the current acquisition time. Of course, this effect becomes pronounced for line separation close to the FWHM of the molecular transition. For instance, when $\nu_S$=200 MHz, the difference in peak intensities is ~0.05 and starts to become a considerable source of measurement uncertainty, albeit it is known *a priori* and can be accounted for in spectroscopic analysis. The FWHM of the line, however, remains unchanged.




**Supplementary Fig. 3** Influence of the collinear (symmetric) dual-comb configuration on the measurement of a Doppler-limited transition.

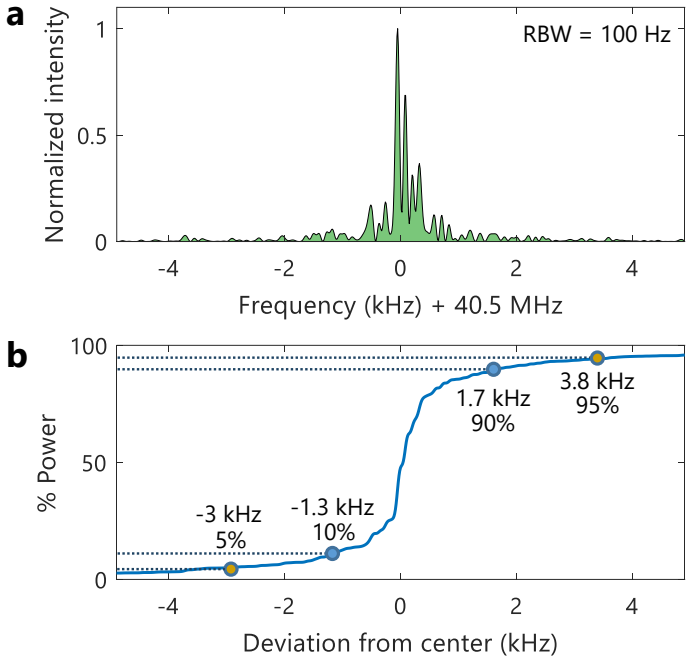
## Supplementary Note 4: Relative comb linewidth

As discussed in the main text, the 3 dB relative comb linewidth with a value below the repetition rate difference may not be a sufficient condition for near-distortion-free computational correction with tooth-resolved RF spectrum. In a typical free-running single-cavity dual-comb laser, the relative comb line has a considerable amount of power outside the half-width range. In our system, we estimated that despite the narrow width of 200 Hz, 80% of the beat note power resides in a 3 kHz bandwidth (see Supplementary Fig. 4), whereas 90% requires almost 7 kHz of frequency span. For the experiment, we picked a repetition rate difference between the two - 4.655 kHz as a reasonable trade-off.

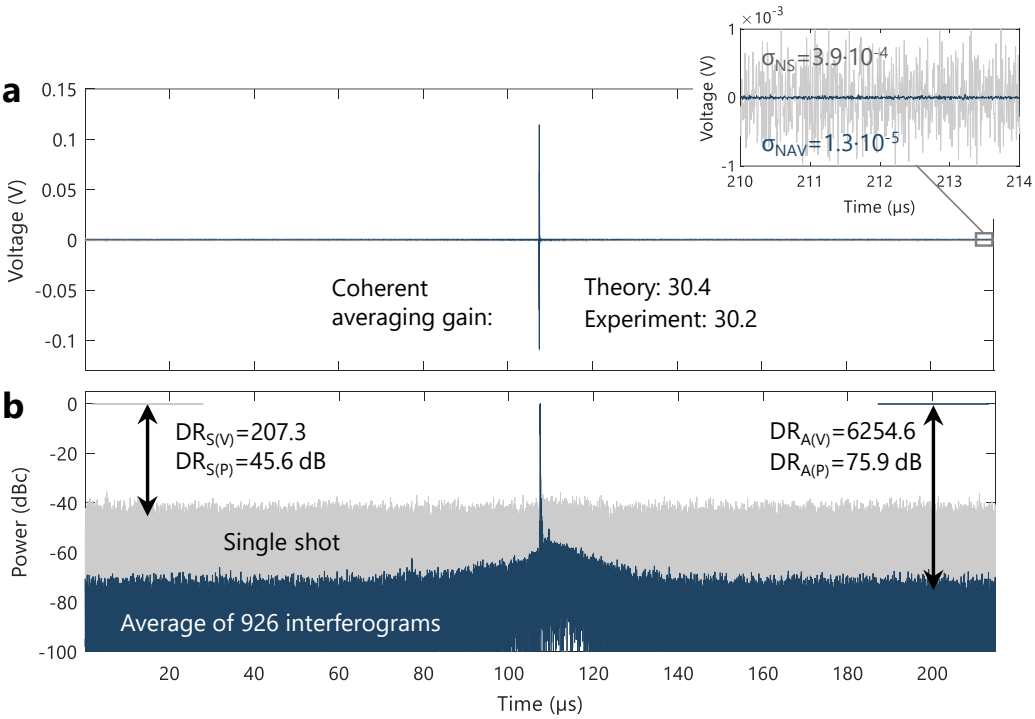
## Supplementary Note 5: Efficacy of the correction averaging routine

To prove an improvement in the dynamic range and effective suppression of amplitude (and phase) noise, we plot a single-shot and 926 coherently averaged interferograms in Supplementary Fig. 5. Panel a) shows a single interferogram frame in a linear voltage scale with an inset illustrating the standard deviation of the noise decreasing by a factor of 30.2 thanks to computational correction. This is 99.3% of the theoretical value for the number of

signals averaged here √926≈30.4. More importantly, despite the large number of averaged traces, the peak value of the trace did not decrease which would be a signature of phase drifts caused by an incorrect retrieval of the correction parameters. Panel b) shows the same signal in a logarithmic scale, where the same improvement factor holds for the ratio of the RMS amplitudes of the interferogram bursts to the RMS level of noise. A closer look around the centerburst of the coherently averaged trace reveals the existence of a periodic interferometric modulation, which was previously buried in noise.




**Supplementary Fig. 4** Amount of power carried in the beat note versus frequency bandwidth. (a) Frequency spectrum in a linear intensity scale. (b) Beat note power plotted against deviation from the center.




**Supplementary Fig. 5** Efficacy of the coherent averaging routine in the time-domain. (a) Single-shot and coherently averaged interferograms. Inset shows the edge of the interferogram trace together with standard deviations of the noise for the single-shot and coherently averaged case. (b) The same plot in a logarithmic scale.

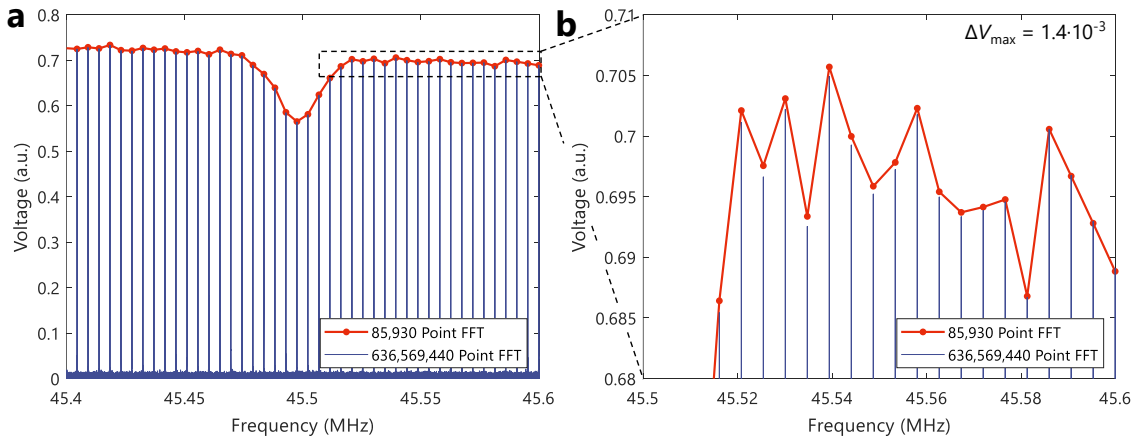
Using the derivation method in[54], we obtain 7.70 bits of dynamic range for the single-shot trace, and 12.61 for the coherently averaged one. If we take into account the variance of a uniformly distributed noise due to quantization[55], the values are 5.90, and 10.82 bits, respectively.

Finally, it would be useful to compare the computational correction with truly phase locked systems. For instance, in Ref.[54], the authors used a pair of phase-locked combs and reached 15.16 bits of dynamic range over 5 seconds, which corresponds to approximately 12.83 bits at 200 ms. The gain of purely-computational coherent averaging is consequently not far (98.2%) from a hardware-based mutual phase lock, however this statement may not hold true over extended time scales on the order of hours or days, as presented in Ref.[55].

## Supplementary Note 6: Calculation of the frequency spectrum

Due to the fact that the interferograms corrected using our routine form a radio frequency harmonic comb, we know the exact locations of all teeth. This allows us to calculate the Fourier Transform spectrum for spectroscopic assessments with a relatively sparse grid and without the necessity of using any windowing, peak amplitude searching algorithms or parabolic interpolation, thus offering huge computational savings and robustness. This strategy is frequently used in DCS[56,57] and ultra-high resolution Fourier Transform Spectroscopy with frequency combs[58].

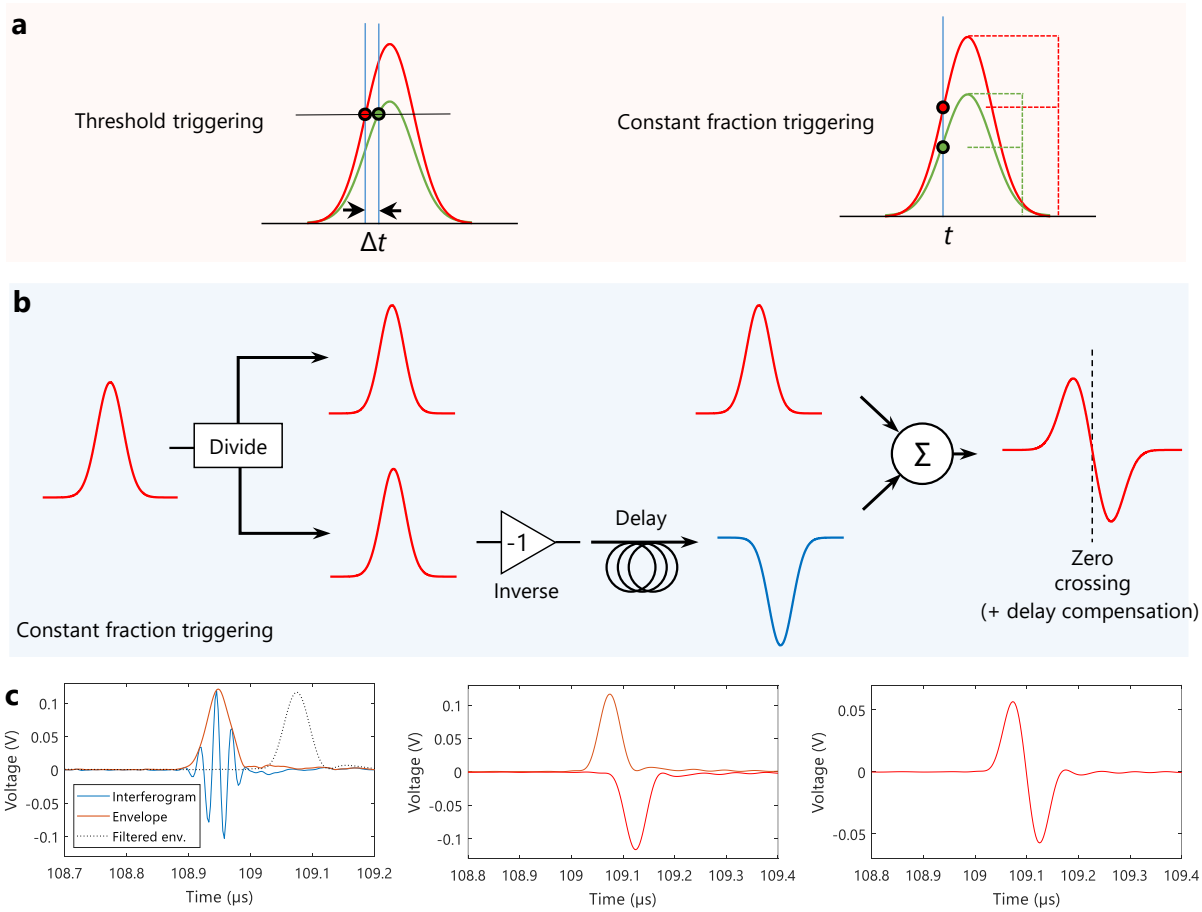
It is well-known that the Discrete Fourier Transform (DFT) can provide a non-distorted amplitude spectrum without spectral leakage of finitely sampled data under the assumption that it is sampled at zero crossings. In other words, the frequencies of the signal must coincide with the sampling bins of the DFT. In the coherently averaged trace, this convenience comes for free – the CEO-free harmonic comb obtained with our algorithm has RF lines located exactly at harmonics of $\Delta f_{rep}$, hence we can obtain a tooth-resolved spectrum just by calculating an $N$-point Fast Fourier Transform without windowing, where $N$ is equal to the number of samples in the coherently averaged interferogram. We compared the two amplitude retrieval techniques, and found maximum deviations in the per mille range probably attributed to numerical errors in the calculation of the ultra-high resolution FFT with such narrow RF comb lines.



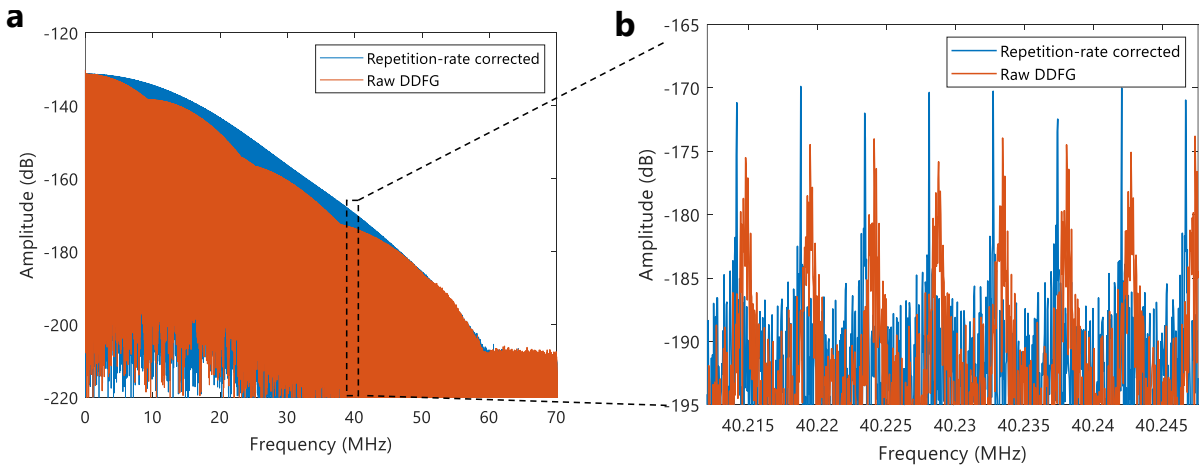

**Supplementary Figure 6.** Comparison of the amplitude spectra calculated with varying sizes of the Discrete Fourier Transform. The red trace corresponds to the case where the size was equal to the number of points in the interferogram, whereas the blue was ~7400 larger. (a**)** One of the absorption lines. (b) Zoomed plot.

## Supplementary Note 7: Correction of variable repetition rate

Supplementary Fig. 7 shows how constant fraction triggering is realized in application to our dual-comb spectrometer. The interferogram is converted into an analytic form, and next its low-pass filtered envelope is used for triggering. Supplementary Fig. 8 shows the efficacy of the repetition rate retrieval and correction illustrating a narrowing of the harmonics of the repetition rate difference.



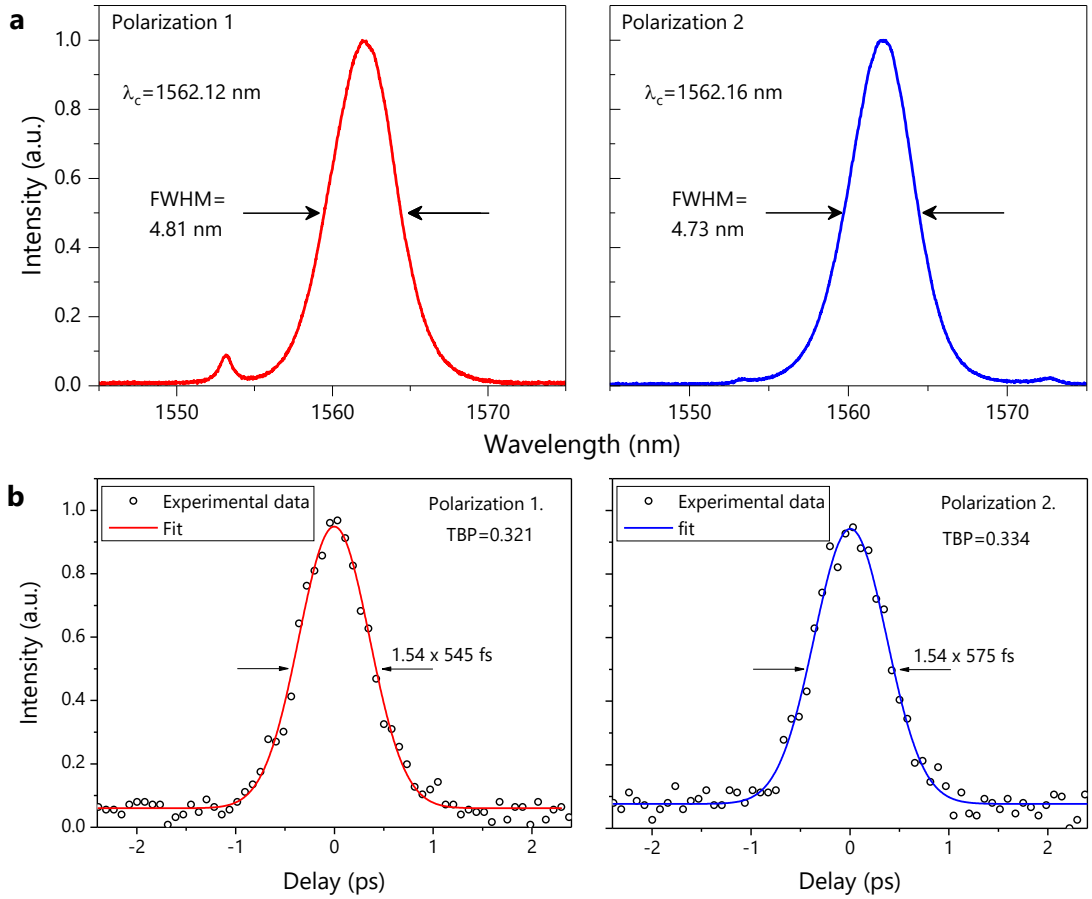

**Supplementary Figure 7.** Operation principle of the constant fraction discriminator. Figure adapted from Ref.[59]. (a) Threshold triggering versus constant fraction triggering. In the first case, a significant time walk is introduced when the peak value changes, which is not the case for constant fraction triggering with a theoretical zero time walk. (b) Schematic diagram of constant fraction triggering. (c) Constant fraction triggering signals obtained in the system. The left panel shows an interferogram after envelope detection and low-pass filtering, the center shows a positive filtered envelope and its time-delayed negative version, which superimposed produce a signal with a well-defined zero crossing on a steep falling edge used for time-of-arrival estimation.




**Supplementary Figure 8.** Frequency spectrum of the digital difference frequency generation (DDFG) signal[39] showing harmonics of the repetition rate difference. (a) After computational resampling, high order harmonics considerably narrow their linewidth and the envelope of the harmonic spectrum takes a more regular shape. (b) Zoomed plot.

## Supplementary Note 8: Pulse width measurement

Supplementary Fig. 9 plots the autocorrelation traces of the two polarizations along with the corresponding optical spectra. The autocorrelation traces were measured using an APE Pulsecheck autocorrelator, while the measured data points were fitted using the $sech^2$ function. The pulse durations of the polarization resolved spectra were 545 fs and 575 fs for polarization 1 and 2, respectively. Taking into account the optical bandwidth (FWHM), the pulses were slightly chirped, which resulted in the time-bandwidth product (TBP) of 0.321 and 0.334 (the theoretical limit is 0.315).




**Supplementary Figure 9.** (a) Optical spectra of the two polarizations measured in a regime similar to that in the main spectroscopy experiment. (b) Corresponding autocorrelation traces.

### Supplementary References


(52) Z. Zhu, K. Ni, Q. Zhou, and G. Wu, "Digital correction method for realizing a phase-stable dual-comb interferometer," Optics Express, vol. 26, no. 13, pp. 16813–16823, 2018.

(53) I. E. Gordon, L. S. Rothman, C. Hill, R. V. Kochanov, Y. Tan, P. F. Bernath, M. Birk, V. Boudon, A. Campargue, K. V. Chance, B. J. Drouin, J.-M. Flaud, R. R. Gamache, J. T. Hodges, D. Jacquemart, V. I. Perevalov, A. Perrin, K. P. Shine, M.-A. H. Smith, J. Tennyson, G. C. Toon, H. Tran, V. G. Tyuterev, A. Barbe, A. G. Császár, V. M. Devi, T. Furtenbacher, J. J. Harrison, J.-M. Hartmann, A. Jolly, T. J. Johnson, T. Karman, I. Kleiner, A. A. Kyuberis, J. Loos, O. M. Lyulin, S. T. Massie, S. N. Mikhailenko, N. Moazzen-Ahmadi, H. S. P. Müller, O. V. Naumenko, A. V. Nikitin, O. L. Polyansky, M. Rey, M. Rotger, S. W. Sharpe, K. Sung, E. Starikova, S. A. Tashkun, J. V. Auwera, G. Wagner, J. Wilzewski, P. Wcisło, S. Yu, and E. J. Zak, "The HITRAN2016 molecular spectroscopic database," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 203, pp. 3–69, 2017.

(54) I. Coddington, W. C. Swann, and N. R. Newbury, "Coherent linear optical sampling at 15 bits of resolution," *Optics letters*, vol. 34, no. 14, pp. 2153–2155, 2009.

(55) J. Roy, J.-D. Deschênes, S. Potvin, and J. Genest, “Continuous real-time correction and averaging for frequency comb interferometry,” *Optics express*, vol. 20, no. 20, pp. 21932–21939, 2012.

(56) E. Baumann, F. R. Giorgetta, W. C. Swann, A. M. Zolot, I. Coddington, and N. R. Newbury, “Spectroscopy of the methane v3 band with an accurate midinfrared coherent dual-comb spectrometer,” *Phys. Rev. A*, vol. 84, no. 6, p. 062513, 2011.

(57) N. Hoghooghi, R. J. Wright, A. S. Makowiecki, W. C. Swann, E. M. Waxman, I. Coddington, and G. B. Rieker, “Broadband coherent cavity-enhanced dual-comb spectroscopy,” Optica, vol. 6, no. 1, p. 28, 2019.

(58) P. Maslowski, K. F. Lee, A. C. Johansson, A. Khodabakhsh, G. Kowzan, L. Rutkowski, A. A. Mills, C. Mohr, J. Jiang, M. E. Fermann, and A. Foltynowicz, “Surpassing the path-limited resolution of Fourier-transform spectrometry with frequency combs,” Phys. Rev. A, vol. 93, no. 2, p. 021802, Feb. 2016 DOI: 10.1103/PhysRevA.93.021802.

(59) “Constant fraction discriminator,” *Wikipedia*. 09-Jun-2014. Available: https://en.wikipedia.org/w/index.php?title=Constant_fraction_discriminator&oldid=612280506. [Accessed: 12-Mar-2019]