SUPPLEMENTARY INFORMATION

1. Ink preparation and recipes

The studied inks were either prepared according to old recipes or purchased from Zecchi (Florence, IT). In Table I, present in the article, we report the compositions of the studied inks and the recipes followed for their preparations. The black inks are iron gall inks, which were prepared in lab following two main recipes containing either oak galls as a source of the gallo-tannic acid (Recipe A in Table I) or preliminarily synthesized gallic acid instead (Recipe B). The recipe A is a historic recipe from Giovanni Alcherio 1411 (on the advice of experts in paper conservation from Opificio delle Pietre Dure Institute in Florence) [12]; the ingredients (water, wine and vinegar) were mixed together with powdered oak galls and left for 1 month to allow for the gallo-tannic acid extraction. The mixture was then heated until the volume was reduced by 1/4. The powdered Arabic gum was added and briefly heated, iron(II) sulfate was added at the end.

2. Samples preparation

Inks were dried in laboratory conditions, blended with polyethylene (PE) powder (Merck), ground, and pressed under a manual hydraulic press at 0.8 GPa to form pellets of 13.2 mm diameter and thickness of about 1 mm. The analytic concentration in the PE pellet was set to be approximately 33 $wt.\%$ that, for most of the studied inks, has revealed to be an optimal concentration, making detectable any eventual features in a relatively wide spectral range. PE was found to be an ideal support for absorption spectroscopy in the THz region thanks to its negligible absorption coefficient (below 1 cm^{-1} see [10] and references therein). In order to study the material's optical properties, wet inks were also deposited on 10 μm thick PE pellicles (IR sample cards, Sigma-Aldrich) and let dry to form films with thickness of tens of μm. Moreover, the iron gall ink (recipe B) was also studied when applied on paper.

3. THz time-domain spectroscopy set-up

Measurements in 0.1 - 4 THz range were performed with a home-made THz-TDS system in transmission configuration. Figure 5 depicts a simplified scheme of our THz-TDS set-up. Optical laser pulses, at $\lambda = 780\ nm$, with a pulse duration of less than 120 fs and repetition rate of 100 MHz (produced by a T-light 780 nm fiber laser from MenloSystems), excites a Low-temperature GaAs photoconductive antenna (PcA) [10], which is biased with sinusoidal voltage of $0 - 30\ Volt$ at a frequency of 10 KHz. The photoexcited carriers are accelerated and shortly after recombine. The abruptly varying photocurrent generates short bursts of electromagnetic radiation with a broad spectrum in the THz region. The emitted THz field is extracted and collected by a hemispherical silicon lens to obtain a divergent beam, which is then collimated and subsequently focused on the sample by a couple of parabolic off-axis mirrors (PMs). The signal transmitted through the sample is again collimated and focused on the detector PcA by a second couple of PMs and an other hemispherical silicon lens optimizes the coupling between the THz field and the dipole of the antenna. A second optical pulse, the probe, generates photo-excited carriers that are accelerated by the THz field, which acts as bias in the detection PcA. The temporal evolution of the photocurrent amplitude in the detection antenna, obtained by changing the time delay between the pump and probe pulses, is directly connected to the electric field amplitude of the THz radiation. This current is amplified by a lock-in amplifier, locked at the bias frequency of the source antenna, and digitalized by an acquisition board. A home-made software acquires the processed signal together with the reading of the delay line encoder and retraces the final time dependent THz field. The working chamber, containing the whole THz set-up, was purged with nitrogen to eliminate the numerous contributions of water vapour, present at the THz frequencies spanned by the experiment. In order to improve the data quality and reduce the effects of external perturbations during the acquisition (e.g. temperature fluctuations) measurements with and without sample are cyclically repeated. The sample is, therefore, mounted on a motorized translation stage to be moved in front of the beam, and for every sample scan a reference scan is also acquired. Each single scan is obtained by a 300 second acquisition at a rate of 10 KHz with a continuous motion of the probe delay line at the velocity of 0.5 mm/sec. Each couple of sample and reference signals are Fourier transformed and their ratio is averaged over all the data, thus giving the

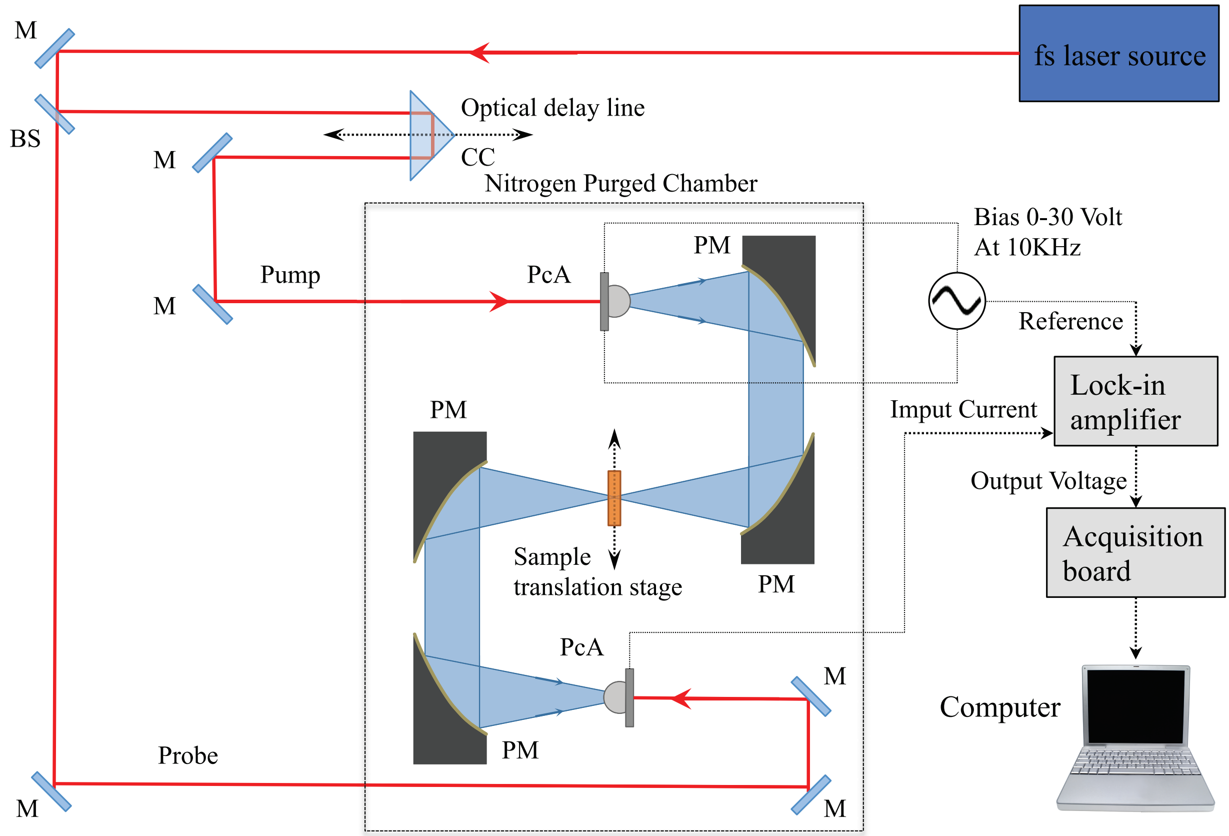

Figure 5. Optical set-up for THz time-domain spectroscopy in transmission configuration: M mirror, BS beam splitter, CC corner cube, PcA photoconductive antenna, PM parabolic mirror.

experimental transfer functions defined in the article.

4. Material parameters extraction from experimental data

In our work we followed the numerical optimisation algorithm proposed by Scheller et al. [14].

The first step is to obtain a rough estimation of n_s and α_s. By neglecting the FP term and the imaginary part of the refractive index in the Fresnel coefficients of eq. 1, we get analytical expressions for the optical parameters [20]:

$$n_s(\omega) = n_0 - \frac{c}{\omega d}\arg\left[H(\omega)\right] \tag{5}$$

$$k_s(\omega) = \frac{c}{\omega d}\left\{ln\left[\frac{4n_0 n_s}{|H(\omega)|(n_0+n_s)^2}\right]\right\} \tag{6}$$

where arg $[H(\omega)]$ is the argument of the complex transfer function. Substituting in H the experimental value H_{exp} and an assumed initial value of d measured with a micrometric screw, we obtain approximated frequency-dependent values of n_s and α_s, which are, moreover, affected by fake oscillations due to the neglected FP effect (see fig. 1 in the article). In order to remove these oscillations, Scheller et al. [14] apply to these raw values of n_s and α_s a band stop filter centred at the FP frequency. We instead implement a polynomial fit of the optical parameters, varying the polynomial order and fitting range as long as the real physical frequency behaviour is revealed and the FP oscillations are removed.

After this first estimation of n_s, α_s, and d, we calculate the full theoretical expression of $H(\omega)$, eq.s 1 and 2 in the article, with the summation of the FP limited to the number of reflections appearing in the time window of the measurement, then we compare it with the experimental one to infer new best values for n_s, α_s, and d. Thus the second step is to minimize the function

$$\triangle H = \sum_{\omega} |H(\omega) - H_{exp}(\omega)| \tag{7}$$

with a numerical optimization on the n_s, α_s for different fixed values of d. We use a Nelder-Mean simplex algorithm with the two scalars ξ and ψ:

$$n_{s,new}(\omega) = \xi\,[n_{s,old}(\omega) - 1] + 1, \tag{8}$$

$$k_{s,new}(\omega) = \psi k_{s,old}(\omega), \tag{9}$$

For every value of d, new values of $n_s(\omega, d)$ and $\alpha_s(\omega, d)$ are calculated by eq.s 5 and 6, filtered, and then optimized minimizing $\triangle H$. Plotting the minima of $\triangle H$ as a function of d we obtain a curve with a minimum in a d_{min} value, which corresponds to the real thickness of the sample. The fitting process is then repeated, starting from the triad, $n_s(\omega, d_{min})$, $\alpha_s(\omega, d_{min})$, and d_{min}, but now with the additional parametrization of d, $d = \zeta d$, in order to refine its value. It is worth to stress that the parametrizations of n_s and α_s through the scalar ξ and ψ do not change their frequency behaviours which are still those inferred from the first step and can be affected by the filtering process. Thus, as third and final step, as already reported by Scheller et all. [14], we perform an optimization of the optical parameters at every frequency step ω_i using the function

$$\triangle H(\omega_i) = |H(\omega_i) - H_{exp}(\omega_i)| \tag{10}$$

The starting values for n_s and α_s are the optimal ones found in the previous step, the parametrizations are the same of eq.s 8 and 9 with the same algorithm, whilst d is always kept fixed to the optimal value estimated before. This last optimization reshapes the frequency features of the optical constants that might have been distorted or erased by the first step evaluation and filtering process. After the polynomial filtering process, this new set of curves of n_s and α_s can be used again as input values of step two of the optimization cycle. Especially for a sample with a short optical path, the optimization must be repeated several times to find the reliable values of the thickness and the optical constants. All the calculations and minimization routines written above by which we analyse all the data reported in this work were performed by executing an in-house developed Matlab code, in Figure 6 we report a block diagram of the fitting procedure.

What described so far concerns the employed analysis for a free standing single slab or layer, thus relevant for our pellet samples. In the case of a bilayer system, as in the layered inks on PE pellicles, the optimization process is similar but starts from a different set of equations. The analysis can be carried out if the optical properties of at least one of the layers and its thickness are known. In our case the PE pellicle has been firstly characterized as a free standing layer by means of the above analysis: we found $d_{PE} = 10\ \mu m$, $n = 1.40$ frequency independent, $\alpha_{PE} \simeq 0$ in the whole probed frequency range. The first step is to consider the bilayer system as a single layer and obtain effective optical parameters using the approximated eq.s 5 and 6 with $d = d_1 + d_2$, where d_1 and d_2 are the thicknesses of the two layers. n_{eff} and α_{eff} can be related to the optical constants of the two layers by simple considerations on the refractive index and absorption coefficient. Let us consider the ink layer as 1 and the PE pellicle as the layer 2, we then get:

$$n_1 = \frac{1}{d_1}\left[(n_{eff} - n0)(d_1 + d_2) - d_2(n2 - n0)\right] + n0, \tag{11}$$

$$k_1 = \frac{1}{d_1} k_{eff}(d_1 + d_2) - k_2 \frac{d_2}{d_1}, \tag{12}$$

these have to be filtered from the FP oscillations following the same procedure described above. Then we can calculate the final values of the parameters following the same procedure of optimization and minimization as in the single slab case simply using the correct expression of the transfer function for a bilayer system, which, for waves at normal incidence,

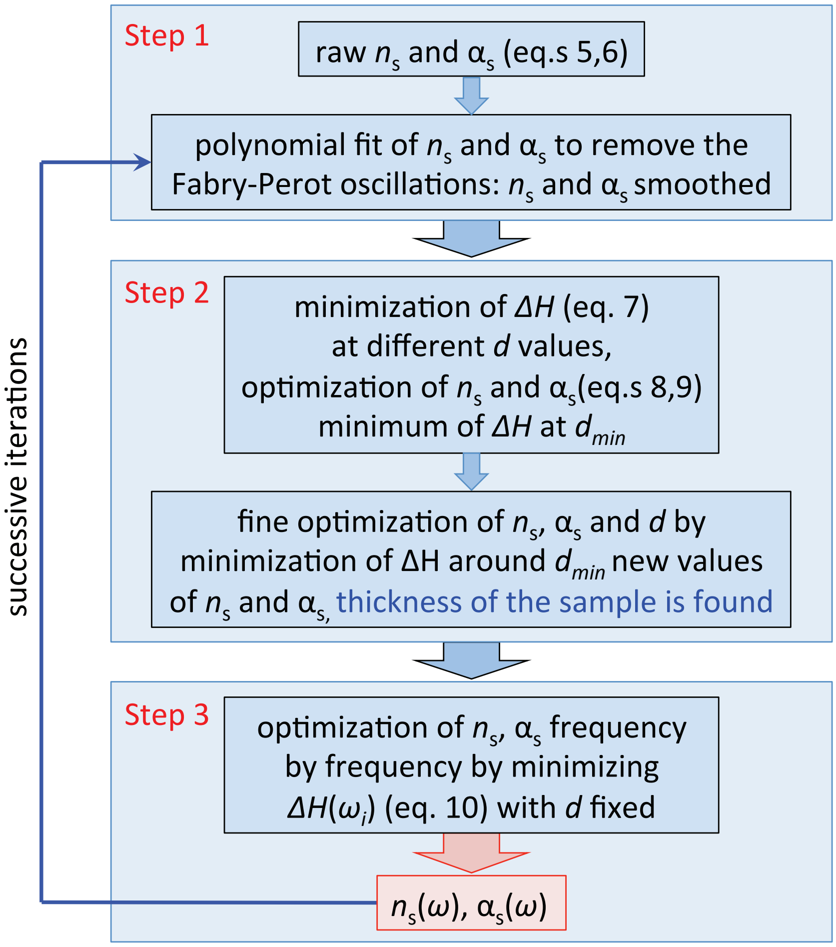

Figure 6. Block diagram of the fitting procedure used for the extraction of the optical parameters, n_S and α_S, and the sample thickness d. The algorithm can be sketched in three main steps: the first one, where a preliminary and approximated evaluation of the refractive index and absorption coefficient is done, the second one, where a minimization routine of the functional ΔH enables the estimation of the correct value of the sample thickness and more reliable values of n_S and α_S, and finally, the third step, by which the real frequency dependency of the optical parameters is obtained through the minimization of ΔH frequency by frequency. With thick samples, where Fabry-Perot reflections are clearly visible and separated, only one cycle from step 1 to step 3 can be enough. For very thin samples successive iterations of the process need to be repeated several times as long as the sample thickness value is stabilized. In each cycle the output parameter of the step 3 are smoothed out by the polynomial fit and used as input parameters again for the step 1.

can be written as [11]:

$$\begin{aligned} H(\omega) &= \frac{E_t(\omega)}{E_i(\omega)} = \\ &= \frac{\tau_{01}\tau_{12}\tau_{20}\ e^{-i\frac{\omega}{c}[d_1\hat{n}_1+d_2\hat{n}_2-n_0(d_1+d_2)]}}{\left[1-\rho_{21}\rho_{20}\ e^{-i\frac{2\omega}{c}d_2\hat{n}_2}\right]\left[1-\rho_{12}\rho_{10}\ e^{-i\frac{2\omega}{c}d_1\hat{n}_1}-\frac{\rho_{20}\rho_{10}\tau_{21}\tau_{12}\ e^{-i\frac{2\omega}{c}(d_1\hat{n}_1+d_2\hat{n}_2)}}{1-\rho_{21}\rho_{20}\ e^{-i\frac{2\omega}{c}d_2\hat{n}_2}}\right]} \end{aligned} \tag{13}$$

where $\hat{n}_i$ are the complex refractive indices, τ_{ij} and ρ_{ij} are the complex transmission and reflection coefficients with $i, j = 0$ for nitrogen, being 1 for ink, and 2 for PE. This expression includes a FP effect with an infinite number of reflexes from the three interfaces. The experimental transfer function, however, is obtained with temporal signal measured in a finite temporal range, so a number of reflexes restricted to this time window should be considered. Taking into account a finite number of reflections is much more complicated and time-consuming. Anyway, considering that the data temporal range is very long compared to the time delays of the reflections in a sample of two very thin layers and that the intensity of each subsequent reflection decays exponentially in time, we expect that the experimental transfer function can be accurately described by eq. 13 [6].