Unveiling The Origin of Spurious Features in THz-TDS of Powder Compacts: Supporting Information

S1: Scattering Angle

For the ray-tracing simulation, it is important to provide a realistic distribution of scattering angles to match the materials under consideration. To achieve this a finite-element frequency domain simulation is modelled in COMSOL which considers a circular Borofloat microsphere within a PTFE matrix. The two materials are modelled with their refractive index values taken at 2 THz. A plane wave is incident upon the microsphere and the scattered electric field is recorded as a function of angle, where 0° corresponds to forward scattering (where the angle of the scattered light is unchanged compared to the incident light). The simulation is run for a range of frequencies from 1 to 3 THz (in 0.5 THz steps) and the square of the electric field is plotted against the scattering angle as shown in Figure S1(a). The plot shows only angles from 0° to 180° as the data is symmetric around 0°. The square of the electric field will be proportional to the probability of a photon scattering into a particular direction. To generate a frequency-agnostic function the maximum of E^2 across all frequencies is calculated at each angle (plotted in the black dashed line) and this is used to generate the customised random distribution function. The customised function "AniRand" (which stands for Anisotropy Randomisation" will return a random angle between ±180° with a probability distribution that follows the shape of E^2 maximum. Figure S1(b) shows a histogram of 10,000 random angles generated by AniRand, overlaid with the shape of E^2 maximum in the red dashed line.

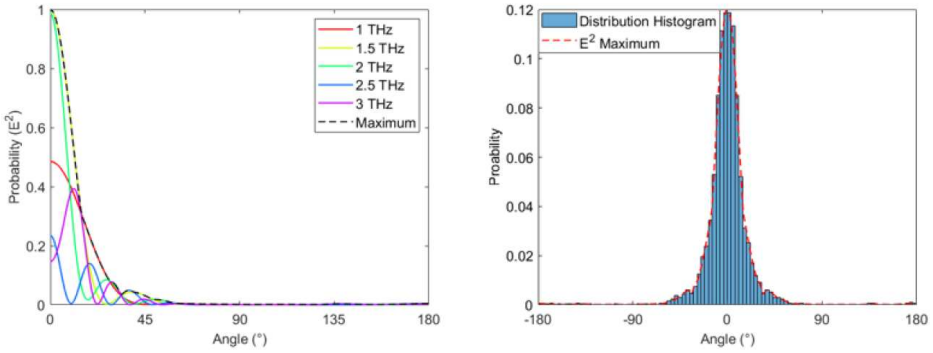

Figure S1: **(a)** Scattering angle probability obtained from numerical simulations. **(b)** A bespoke randomised number distribution is produced in Matlab which follows the distribution of the scattering angles.

S2: Sample Size

The real-world samples consist of a 1 mm depth slab of material, illuminated by a focused THz beam with a 3 mm diameter focal spot. The number of Borofloat microspheres within the illuminated region is calculated by assuming a cylindrical volume of material with diameter 3 mm and depth 1 mm. Within this volume, the total number of Borofloat microspheres N can be calculated as a function of microsphere radius R and concentration ρ shown in Figure S2(a).

The cylindrical volume of the THz focus (V_f) is 7.07 mm^3, based on a focal spot radius (r) of 1.5 mm, and a sample depth (d) of 1 mm. Within this volume the number of Borofloat microspheres (N_{3D}) is shown below, where ρ is the fractional concentration of microspheres, and V_b is the volume of the microspheres.

$$N_{3D} = \frac{V_f\ \rho}{V_b} = \frac{\rho\ \pi\ r^2\ d}{\frac{4}{3}\ \pi\ R^3} \tag{1}$$

For the case of a 2D simulated sample, the sample width, W, must be of correct size that the number of microspheres in the simulation matches the number of microspheres for the 3D scenario. The number of microspheres in the 2D sample (N_{2D}) is defined by the sample area A_f, the fractional concentration (ρ), and the microsphere area (A_b). The sample area A_f is a product of the sample depth (1 mm) and the sample width, W.

$$N_{2D} = \frac{A_f\ \rho}{A_b} = \frac{\rho\ W\ d}{\pi\ R^2} \tag{2}$$

By equating the values of N_{3D} and N_{2D}, the concentration terms ρ cancel out and the sample width, W, can be found as a function of the microsphere radius.

$$W = \frac{\pi\ r^2}{\frac{4}{3}\ R} \tag{3}$$

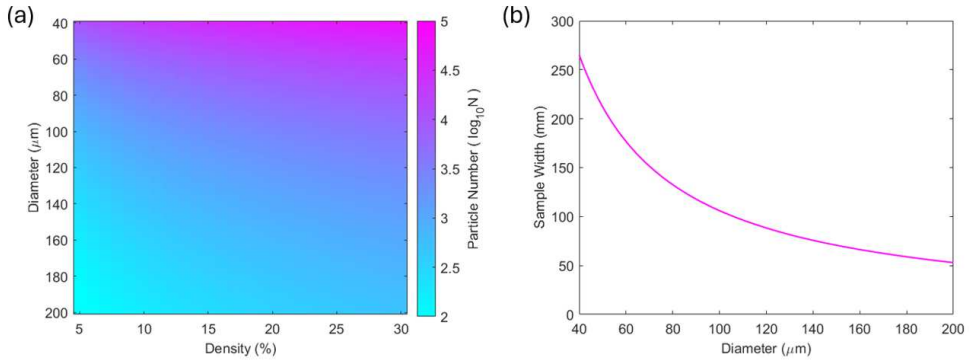

Figure S2: **(a)** Numerical calculations are used to determine the number of Borofloat microspheres that would be expected within a 3D measurement volume (depending on radius and density). **(b)** The width, W, of a 2D sample is tailored to contain a similar number of microspheres.

S3: Ray Tracing Simulation

Geometry Creation

1. Firstly the diameter D of the Borofloat microspheres and their area concentration ρ within the PTFE matrix is defined. A typical value here is 98 μm diameter microspheres with 20 % area concentration.

2. Based on the microsphere diameter the width, W, of the slab is determined following the procedure in supplementary section S2. This step is performed to ensure that a 2D sample contains a similar number of microspheres as the 3D scenario from the experiments. The height H of the slab is 1 mm.

3. A pixelated slab (an empty matrix M) is generated with pixel size 5 μm, depth 1 mm, and width W.

4. The target number N of microspheres is calculated based on the concentration ρ, microsphere diameter D, and slab area.

5. A 2D matrix ($2xN$) of random numbers is generated which represents the centre points of the Borofloat microspheres in x and y coordinates. Random numbers are generated 0 - 1 following a uniform distribution, then multiplied by the size of the slab in the X and Y dimensions. However, instead of allowing locations for example in the X axis from 0 to W, the limits of the X and Y dimensions here are modified to account for the diameter D of the microspheres to prevent a microsphere being partly outside the edges of the slab. Therefore the limits are modified to $\frac{D}{2} \leq x \leq W - \frac{D}{2}$, and $\frac{D}{2} \leq y \leq H - \frac{D}{2}$.

6. To prevent the un-physical scenario of microspheres spatially overlapping, the distance between each microsphere centre-point and its nearest neighbour is calculated to find any centre-point combinations which are closer than D using the Matlab function 'pdist2'. When any such elements are found, one of the pair is chosen to be redistributed to a new randomised location.

7. The previous step is looped repeatedly until the microspheres stop moving, which determines that all microspheres are then non-overlapping and randomly distributed. For concentrations $\leq$ 30 % this method works efficiently and requires only a few iterations to correctly distribute the microspheres.

8. Using the array of microsphere centre locations the matrix M is then filled in a binary nature, where pixels of distance from each centre point $\leq \frac{D}{2}$ are filled with 1's (representing Borofloat material), and all other pixels keep a value of 0 (representing PTFE).

Ray-Tracing

1. A ray starts at the upper surface of the slab, with a randomised location x in the $\hat{x}$ axis, and a randomised incident angle θ normal to the surface. To prevent the ray from starting too close to the horizontal edge and hitting the side of the slab, the x values are limited to a uniform distribution between $\frac{1}{4}W \leq x \leq \frac{3}{4}W$. The incident angles are uniformly distributed between $\pm$ 20° to provide a range of incident angles. The y location at the upper surface of the slab is defined as 0. The

ray is assigned an initial intensity $I = 1$, which will be used to model the effects of absorption while passing through the slab.

2. A line is generated using the parameters (x, y, θ) which provides the direction of propagation in an analytical formulation. The line is converted to a list of pixel locations that it passes through.

3. A ray then starts to move along this path, pixel by pixel. At each step, the time taken to traverse the pixel is calculated using the pixel dimension (5 μm) divided by the speed (c/n), and added to a counter which will be used to sum the total time-of-flight for the ray. The intensity I at each step is reduced according to the imaginary refractive index k and the distance traveled (equal to the pixel size).

4. When the ray moves into a new pixel, if it detects that it has entered a Borofloat microsphere (if $M_{i,j} = 1$) then the angle of propagation θ is altered according to the randomised distribution function AniRand. The propagation angle is now 'locked' while the ray is inside the microsphere which means the ray angle cannot change again; the lock only will be released once the ray exits the microsphere. AniRand is a customised function that returns a randomised angle between $\pm$ 180° with a specific probability distribution (discussed in detail in supplementary section S1). The AniRand function was obtained by modelling in COMSOL the angular distribution of scattered intensity for a 200 μm diameter microsphere at a range of frequencies. This function is then utilised in the ray tracing model for all microsphere sizes.

5. In the case of very large scattering angles it is possible the ray will immediately leave the microsphere, and therefore will be back-scattered. In this scenario the ray will incur a time increase of 30 fs as the simulation will detect the ray was present inside a single pixel of the microsphere (at pixel size of 5 μm).

6. The ray will continue to move through the microsphere along its newly defined line, with the time-of-flight through each pixel now taking slightly longer owing to the increased real refractive index of Borofloat compared to PTFE.

7. When the ray exits the Borofloat microsphere it continues to travel at the same angle. The lock on the angle is now removed, ready for a new interaction with a microsphere if it occurs.

8. Steps 4, 5, 6, and 7 are repeated until a ray is detected to have reached a boundary of the slab. If the ray exits the slab on the top or side surfaces then its existence is ignored. If the ray exits on the lower surface of the slab then its total time-of-flight and intensity are recorded.

Data Processing

1. A total of 1 million rays are traced through the sample as this is shown to provide a stable result.

2. The time-of-flight of each ray is binned, utilising 600 equally spaced bins between 0 and 20 ps, where the intensity of each ray determines its contributing weight in the bin, thereby accounting for absorption effects. The binned data gives the temporal profile of rays traced through the sample.

3. A Fourier transform is used to convert the temporal profile into the frequency domain.