Table of contents

1. Data content

Provided dataset is divided into three subsets:

- Laser Scanning Microscopy (LSM) measurements, representing LSM images of 9k7 samples cross-sections and surface and stored in `lext` format in the folder “raw data”
- Pores data, which contains pores coordinates, pores depths, equivalent radii of pores, lengths of major and minor axes of equivalent ellipse and shape factor of equivalent ellipse.
- Surface roughness data, that contains calculated Power Spectral Densities (PSD) functions for surface roughness: experimental as well as theoretical one.

For more details see description.txt file in corresponding data folder.

1.1. Laser Scanning Microscopy measurements

The folder “00_raw data” contains the captured data of the cross-sectional surface in the folder “cross section. Here, the data are organized by sample name (9k7-001.lext … 9k7-010.lext)

The folder “planar” contains the data of the in-polishing plane measurements, organized in subfolders according to the respective polishing time. For each polishing step, 5 measurement points are considered. Figure 1 illustrates the measurement positions and shows their naming.

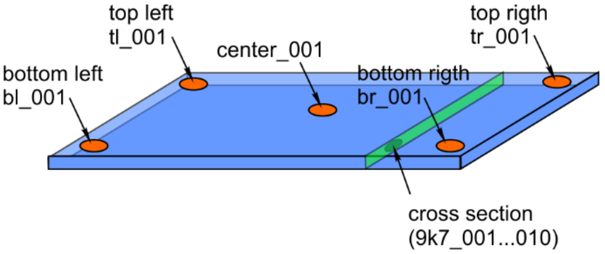

Figure 1: measurement positions.

1.2. Pores data

The folder “01_pores data” follows the same structure, containing the subfolders “cross section” and “planar” with the respective pores characteristics of the cross section and in polishing plane.

1.3. Surface roughness data

The folder “02_surface roughness data” contains PSD functions obtained from the experimental data (*.lext) and theoretical PSD functions used to approximate experimental function. Experimental PSD functions are for center point and 0 minutes of polishing and for edge point and 70 minutes of polishing.

2. Preprocessing

Before the data analysis the image pre-processing is required. This procedure could include the following steps depending on specific problem:

1) Removal of the surface form. Form is the low frequency component of surface texture in contrast to waviness (medium frequencies) and roughness (high frequencies). The option “Data process” → “Level”→ “Polynomial background” of Gwyddion software [1] is used to perform this operation.
2) Spike noise removal from the map. Spikes are non-physically high or low values on the intensity map. This noise is caused by incorrect detection of reflection intensity from areas with high height difference, edges etc. The Gwyddion software [1] is used for this purpose (“Data process” → “Correct Data”→ “Mask of Disconnected”). However, spike removal distorts the local height map and therefore could affect the number of pores that are segmented. This could lead to decrease of shallow pores with small radii. Data provided for the pores characteristics obtained without the spike removal. We compared pores distributions with and without spikes removal and obtained that pores ensambles statistics remain practically the same, but number of pores with small radii ($R < 0.5$ μm) decrease by ~ 20%. Therefore, the spike removal is conditioned by the balance between the gain in noise removal and the loss due to the distortion of the height map.

3. Pores masking

The LSM dataset represented as images (an intensity map and a height map) is used for pores identification. The darkness level corresponds to the intensity level of reflected beam or height value, of a given pixel of the image respectively. Determination of all pixels belonging to each pore gives a set of spots covering (masking) pores on the map. Both images (maps) can be used to define a mask highlighting the pores. The choice of the image in a particular case is determined by the contrast between the pores and the rest of the surface, which may be better for this or that image, depending on the features of the surface structure.

Surface masking in application to the pores segmentation consists of several steps: (i) surface segmentation, (ii) establishing a zero surface level; (iii) removing shallow spots.

3.1. Segmentation

The pores are identified by the technique implemented using Gwyddion software [1]. The intensity or height map obtained by LSM is covered by a mask using different algorithms: watershed-type, thresholding [2] to mark all the notable recesses on the surface. The option “Data process” → “Grains”→ “Mark by Segmentation” of Gwyddion software' is used for this purpose. The most optimal method of segmentation applied to a series of samples obtained under the same conditions is the one in which the maximum number of segmentation parameters is fixed. The fixed parameters are related to the average surface characteristics, while the varying parameters take into account the presence of heterogeneity (noise) inherent to these surface characteristics.

3.2. Surface zero level

After the initial surface segmentation, the surface zero level is set according to the zero level of the unmasked surface. ("Data process" → "Level"→ "Zero Mean Value" with option "Exclude masked region"). The mask obtained in this way may include spots corresponding to pores as well as spots corresponding to recesses associated with surface roughness. To exclude the latter from consideration, the separation on the depth of spots was done. Depth of the spot was calculated using a height map as the distance between surface zero level and deepest point of the surface under the spot.

3.3. Removing shallow spots

The spots with depths less than threshold value were removed from the mask. Value of the threshold could be chosen based on peculiarities of the surface, in our case it was double root mean square (RMS) roughness value, calculated for the unmasked surface. The mask obtained by this procedure consists of spots covering all visible pores except the shallowest ones, which are difficult to distinguish from surface roughness.

Described procedure allows determining the local minima (dark regions) and delineating the valleys that surround them. The obtained topographical data, divided into two categories, namely the pores and the surrounding surface, can be further subjected to a variety of analyses.

3.4. Data masking on cross-section scans (Pore identification)

*The calculations based on the raw data in folder "**Laser Scanning Microscopy measurements/Cross-sections**"*

The identification of cross-sectional pores was carried out by the procedure described above for the LSM images of six different cross sections (256 x 256 μm^2) of the 9k7 LTCC sample in the folder "cross-sections". The intensity map was used for the identification. The following set of parameters of Segmentation procedure was found appropriate and used for pores masking:

Gaussian smoothing	Add gradient	Add curvature	Prefill level	Prefill from minima
5.61	31.84	0	7.96	10.83

Only one parameter "Barrier level" was varied from sample to sample in the range 19.54 – 40.10 to ensure correct pores filling. An example of pore identification is shown in Figure 4, which shows part of the surface of one of the cross sections.

3.5. Data masking on in-plane surface scans (Pore identification)

*The calculations based on the raw data in folder "**00_raw data**"*

The identification of surface pores was carried out by the same procedure as for in-plane ones. The LSM images of the 9k7 LTCC samples without polishing and polished for 20, 50 and 70 minutes were used for the identification. The LSM data can be found in folders "0 min polishing", "20 min polishing", "50 min polishing" and "70 min polishing", respectively.

The difference between these samples and samples with cross-sections is that the former have much higher roughness. The masking procedure for the surface should take into account raising structures with a height up to several hundred nanometers, which are leveled during the polishing process. Therefore, only sufficiently deep pores, the depth of which exceeds the level of surface roughness, can be reliably identified. As a result, the set of pores identified on the surface is much smaller than that on the cross sections.

Masking of the surface pores was performed using a thresholding procedure, since the surface is very rough and water-shed algorithms couldn't provide needed masking quality.

4. Pore analysis

*Pores characteristics stored in folder "**Pores data**"*

4.1. Pores size

Pore sizes are characterized by the pore radius. Since pores are not spherical, for the size characterization the equivalent radius is used. Equivalent radius determines the pore cross-sectional area A_{CSp} and is defined as a radius of a circle (equivalent circle) with area:

$$A_{eq} = \pi R^2_{eq} \qquad (1)$$

equal to the area of the pore cross section

$$A_{eq} = A_{CSp} \qquad (2)$$

4.2. Shape factor

The pores shapes are characterized by the equivalent ellipse. Equivalent ellipse is the ellipse which has the same second order moments (distribution of area/mass) in the horizontal cross-sectional plane as the pore. Pore shape factor η is calculated as relation between major L_M and minor axes l_m of equivalent ellipse:

$$\eta = L_M/l_m. \qquad (3)$$

4.3. Pores inhomogeneity

The program for the pores inhomogeneity factor calculation is available at [3].

5. Polishing analysis

*The calculations are based on the raw data in folder "**Laser Scanning Microscopy measurements/Cross-sections**". Resulting data are stored in "**Surface roughness data**"*

5.1. Power spectral density (PSD) calculation

The spectral power density (PSD) function S represents the height distribution of surface structures as a function of spatial frequencies k, which is the reciprocal of the lateral size of the surface structure. The 1D PSD function S can be calculated from surface height data $z(x)$, where x is a coordinate, and k is a spatial frequency [4]:

$$S = \frac{1}{L}\left|\int_0^L z(x)e^{-2\pi ikx}dx\right| \qquad (4)$$

L is the size of the scan. The unit of S is [μm^3]. For the analysis of PSD function in different spatial regions the experimental PSD function, S_{exp} is decomposed on the sum of theoretical (analytical) PSD functions S_{th}, $S_{exp} = \sum S_{th}$ [4]. Experimental PSD is approximated with good accuracy by the sum of two analytical PSD functions, each described by the ABC-model, which is extensively used in PSD analysis [5]. Table 1 summarizes the fitting constants.

$$S_{ABC} = \frac{A}{\left(1+(Bf_x)^2\right)^{C/2}} \qquad (5)$$

Table 1: Fitting coefficients of the model PSD functions found to be suitable for modeling of experimental data

	center, 0 min polishing		edge, 70 min polishing	
	ABC1	ABC2	ABC1	ABC2
A[μm^3]	0.0094	3.9	0.0019	1.9
B[μm]	0.78	10.5	0.941	11.01
C	5	5	4	4

The PSD function is a suitable tool to extract different types of surface deviations, e.g. roughness and waviness, from the surface topography. By analyzing them, it is possible to assess which structural quantities occur frequently and whether periodic structures are present on the surface.

5.2. RMS height for roughness and waviness

Cut-off wavelength for roughness: 1 μm
Cut-off wavelength for waviness: 3 μm
The sequence of operations used in this technique to characterize the surface of a porous material is as follows: (i) identification of pores from a height or intensity map; separation of the surface occupied by the pores and the surface between the pores; (ii) calculation of the spatial heterogeneity of the pore distribution; (iii) calculation of material porosity and dimensional characteristics of a pore ensemble: distributions of equivalent radius and shape factor; (iv) calculation of height distribution and RMS height for the surface in pores and outside the pores; calculation of the inclination angle distribution; (v) building a PSD function and determining cutoff lengths to separate (distinguish) roughness and waviness; (vi) application of zero-order Gaussian regression filter (ISO 16610-71) to height data using the calculated cutoff lengths to calculate RMS roughness and RMS waviness. 6. References

References

[1] Nečas D and Klapetek P 2022 *GWYDDION: Scanning probe microscopy data visualisation and analysis* (Czech Metrology Institute, Department of Nanometrology)

[2] Klapetek P, Ohlidal I, Franta D, Montaigne-Ramil A, Bonnani A, Stifter D and Sitter H 2003 *Acta Physica Slovaca* **53** 223–30

[3] Lubov M and Kharlamov V 2022 *Pores inhomogeineity measure, https://github.com/modellingioffe/pores_inhomogeneity*

[4] Senthilkumar M, Sahoo N K, Thakur S and Tokas R B 2005 *Applied Surface Science* **252** 1608–19

[5] Church E L, Takacs P Z and Leonard T A 1989 The Prediction Of BRDFs From Surface Profile Measurements *Scatter from Optical Components 33rd Annual Techincal Symposium (San Diego, Monday 7 August 1989) (SPIE Proceedings)* ed J C Stover (SPIE) p 136