Supplementary Material for

R-VINE COPULAS FOR DATA-DRIVEN QUANTIFICATION OF DESCRIPTOR RELATIONSHIPS IN POROUS MATERIALS

MATTHIAS NEUMANN[1,⋆], PHILLIP GRÄFENSTEINER[1,⋆], EDUARDO MACHADO CHARRY[2,3], ULRICH HIRN[3,4], ANDRÉ HILGER[5], INGO MANKE[5], ROBERT SCHENNACH[2,3], VOLKER SCHMIDT[1], KARIN ZOJER[2,3]

[1] *Institute of Stochastics, Ulm University, Helmholtzstraße 18, 89069 Ulm, Germany*

[2] *Institute of Solid State Physics, NAWI Graz, Graz University of Technology, Petersgasse 16/II, 8010 Graz, Austria*

[3] *Christian Doppler Laboratory for Mass Transport through Paper, Graz University of Technology, Petersgasse 16/II, 8010 Graz, Austria*

[4] *Institute of Bioproducts and Paper Technology, Graz University of Technology, Inffeldgasse 23, 8010 Graz, Austria*

[5] *Institute of Applied Materials, Helmholtz-Zentrum Berlin für Materialien und Energie, Hahn-Meitner-Platz 1, 14109 Berlin, Germany*

[⋆] *M..N and P.G. contributed equally to this work*

S1. Decomposition of multivariate probability densities. Here we briefly summarize the mathematical arguments needed to arrive at a decomposition of a multivariate probability density into a product of univariate probability densities and bivariate copula densities, see *e.g.* [3, 4] for further details. First, the necessary steps are provided for the decomposition of a general multivariate density function $f_{1,\ldots,n} : \mathbb{R}^n \to [0, \infty)$ for any fixed integer $n \in \{2, 3, \ldots\}$. Afterwards, the example of a quadrivariate density $f_{1,2,3,4} : \mathbb{R}^4 \to [0, \infty)$ will be used to illustrate these steps.

S1.1. General case. Suppose that we are given n random variables $W_1, \ldots, W_n$ with absolutely continuous joint distribution and differentiable multivariate distribution function. While in the main text we simply denoted their n-variate joint density of $(W_1, \ldots, W_n)$ by f, we will need the additional specification of the indices $1, \ldots, n$ here, *i.e.*, instead of f we will write $f_{1,\ldots,n}$ in the following. The ultimate goal is to represent $f_{1,\ldots,n}$ by the univariate densities f_i of W_i and some (conditional) bivariate copula densities $\widetilde{c}_{ij}$ evaluated by means of some conditional distribution functions $\widetilde{F}_i$ and $\widetilde{F}_j$ of W_i and W_j, respectively, where $i, j \in \{1, \ldots, n\}$. To get there, three steps are required: (i) Rewrite the n-variate density $f_{1,\ldots,n}$ in terms of univariate conditional densities using the chain rule for conditional probability densities, and (ii) consecutively convert the univariate conditional densities into bivariate conditional densities. Finally, in step (iii), each bivariate density is expressed by a bivariate copula density function. The procedure stops once each of the bivariate conditional densities is written as a product of marginal densities and bivariate copula density functions.

Rewriting $f_{1,\ldots,n}$ as product of univariate conditional densities. To express the n-variate density $f_{1,\ldots,n}$ in terms of conditional densities, the chain rule for conditional probability densities is used. Intuitively, this corresponds to consecutive drawing of the sample $(w_1, \ldots, w_n)$ from $f_{1,\ldots,n}$, each time conditioning on the already drawn values. The resulting identity then reads as

$$\begin{aligned} f_{1,\ldots,n}(w_1, \ldots, w_n) &= f_{n;1\ldots n-1}(w_n; w_1, \ldots, w_{n-1}) f_{n-1;1\ldots n-2}(w_{n-1}; w_1, \ldots, w_{n-2}) \\ &\times \ldots \times f_{2;1}(w_2; w_1)\, f_1(w_1), \qquad \text{for all } w_1, \ldots, w_n \in \mathbb{R}, \end{aligned} \tag{S1}$$

where $f_{i;1\ldots i-1}(\cdot; w_1, \ldots, w_{i-1}) : \mathbb{R} \to [0, \infty)$ denotes the univariate conditional density of W_i given that $(W_1, \ldots, W_{i-1}) = (w_1, \ldots, w_{i-1})$, for all $i = 2, \ldots, n$. Note that this decomposition is possible regardless of the order of the variables with respect to which we start the draw. In the formulation considered above, the variables are chosen in ascending order, as it drastically simplifies the notation in the following arguments.

Converting univariate conditional densities into bivariate conditional densities. Note that the univariate conditional densities on the right-hand side of Eq. (S1) are defined by

$$f_{i;1\ldots i-1}(w_i; w_1, \ldots, w_{i-1}) = \frac{f_{1,\ldots,i}(w_1, \ldots, w_i)}{f_{1,\ldots,i-1}(w_1, \ldots, w_{i-1})}, \tag{S2}$$

for all $w_1, \ldots, w_i \in \mathbb{R}$ and $i = 2, \ldots, n$. For $i = 2$ the numerator on the right-hand side of Eq. (S2) is already a bivariate density. This is desired, as we can later apply Eq. (S6) to the numerator, which cancels out the denominator in Eq. (S2). If $i \geqslant 3$, an additional step is necessary, where we use the fact that the conditional bivariate density $f_{1,i;2,\ldots,i-1}$ of (W_1, W_i), given that $(W_2, \ldots, W_{i-1}) = (w_2, \ldots, w_{i-1})$, is defined by

$$f_{1,i;2,\ldots,i-1}(w_1, w_i; w_2, \ldots, w_{i-1}) = \frac{f_{1,\ldots,i}(w_1, \ldots, w_i)}{f_{2,\ldots,i-1}(w_2, \ldots, w_{i-1})}, \quad \text{for all } w_i, \ldots, w_j \in \mathbb{R}. \tag{S3}$$

Thus, it holds that

$$f_{1,\ldots,i}(w_1, \ldots, w_i) = f_{1,i;2,\ldots,i-1}(w_1, w_i; w_2, \ldots, w_{i-1}) f_{2,\ldots,i-1}(w_2, \ldots, w_{i-1}), \tag{S4}$$

for all $w_1, \ldots, w_i \in \mathbb{R}$ and $i \geqslant 3$. Inserting Eq. (S4) into the numerator of Eq. (S2) yields another representation of $f_{i;1\ldots i-1}$ for $i \geqslant 3$ through

$$\begin{aligned} f_{i;1\ldots i-1}(w_i; w_1, \ldots, w_{i-1}) &= \frac{f_{1,i;2,\ldots,i-1}(w_1, w_i; w_2, \ldots, w_{i-1}) f_{2,\ldots,i-1}(w_2, \ldots, w_{i-1})}{f_{1,\ldots,i-1}(w_1, \ldots, w_{i-1})}, \\ &= \frac{f_{1,i;2\ldots i-1}(w_1, w_i; w_2, \ldots, w_{i-1})}{f_{1;2\ldots i-1}(w_1; w_2, \ldots, w_{i-1})} \quad \text{for all } w_1, \ldots, w_i \in \mathbb{R}, \end{aligned} \tag{S5}$$

in which the numerator is now again a bivariate density. Analogously to Eq. (S2) for the case $i = 2$, we will later apply Eq. (S8) to the numerator of Eq. (S5), which again cancels out the denominator of Eq. (S5).

Expressing arbitrary bivariate densities by bivariate copula densities. Sklar's theorem [3, 4] states that any bivariate probability density $f_{i,j} : \mathbb{R}^2 \to [0, \infty)$ of a two-dimensional random vector (W_i, W_j) can be expressed by a bivariate copula density $c_{i,j} : [0,1]^2 \to [0,\infty)$ via

$$f_{i,j}(w_i, w_j) = c_{i,j}\big(F_i(w_i), F_j(w_j)\big)\; f_i(w_i)\; f_j(w_j), \quad \text{for all } w_i, w_j \in \mathbb{R}, \tag{S6}$$

where F_i and F_j are the (univariate) cumulative distribution functions of W_i and W_j, respectively, and $c_{i,j}$ is the bivariate copula density of some copula function $C_{i,j} : [0,1]^2 \to [0,1]$, that is

$$c_{i,j}(u_i, u_j) = \frac{\partial^2 C_{i,j}(u_i, u_j)}{\partial u_i \partial u_j}, \qquad \text{for all } u_i, u_j \in [0,1]. \tag{S7}$$

Bivariate copula functions C_{ij} can be chosen from a wide range of parametric function families to fit the bivariate density $f_{i,j}$ best. For an overview on such function families, the reader may refer to, e.g., the selection given in Tables S2 and S3 below, see also [1, 2]. By construction, copula densities couple the marginal distributions of W_i and W_j (via F_i and F_j) to a bivariate (joint) probability density, by means of one or several additional copula-specific parameters, that contain the relation strength between the two variables W_i and W_j, alike the correlation coefficient in bivariate Gaussian distributions.

In the case of a conditional bivariate density of the form $f_{i,j;i+1,\ldots,j-1}$ with $i+1 < j$, using again Sklar's theorem, a representation analogous to Eq. (S6) is acquired by adding the indices $i+1, \ldots, j-1$ of the conditioning variables to every factor, yielding

$$f_{i,j;i+1,\ldots,j-1} = c_{i,j;i+1,\ldots,j-1}\big(F_{i;i+1,\ldots,j-1}, F_{j;i+1,\ldots,j-1}\big)\; f_{i;i+1,\ldots,j-1}\; f_{j;i+1,\ldots,j-1}, \tag{S8}$$

where we suppress the arguments of the functions for better readability. At this point, the simplifying assumption is used, *i.e.*, we assume that the copula density $c_{i,j;i+1,\ldots,j-1}$ in Eq. (S8) does not depend on the specific values of $(w_{i+1}, \ldots, w_{j-1})$. This allows us to represent the numerator in Eq. (S5) through a (conditional) bivariate copula density.

Assembling the multivariate probability density $f_{1,\ldots,n}$. We can now express each conditional density in Eq. (S1) by a bivariate conditional density via Eq. (S5). Subsequently, each of these bivariate densities is expressed by a bivariate copula density as prescribed in Eqs. (S6) and (S8. This yields the following expression

$$f_{1,\ldots,n} = \prod_{i=1}^{n-1} \prod_{j=i+1}^{n-1} c_{i,j;i+1,\ldots,j-1}\big(F_{i;i+1\ldots,j-1}, F_{j;i+1\ldots,j-1}\big) \prod_{k=1}^{n} f_k, \tag{S9}$$

where we again suppress the arguments of functions for better readability.

S1.2. Example: Decomposition of a quadrivariate probasbility density. For the case $n = 4$, we now demonstrate all the steps considered above, necessary to decompose the joint (quadrivariate) probability density $f_{1,2,3,4} : \mathbb{R}^4 \to [0,\infty)$ of a 4-dimensional random vector (W_1, W_2, W_3, W_4) into a product of univariate (marginal) probability densities and bivariate copula densities. To this end, it is most instructive to show this by adding the random variables W_1, W_2, W_3, W_4 one by one.

If we consider only one variable W_1, then f_1 is its marginal density and no decomposition is necessary. With one more additional variable, W_2, we need to decompose the joint desnity $f_{1,2}$ of (W_1, W_2). Following Eq. (S6), we readily get that

$$f_{1,2}(w_1, w_2) = c_{1,2}\big(F_1(w_1), F_2(w_2)\big)\; f_1(w_1)\; f_2(w_2), \quad \text{for all } w_1, w_2 \in \mathbb{R}. \tag{S10}$$

Adding a third variable, the dependencies between the variables W_1, W_2, W_3 in $f_{1,2,3}$ cannot be captured by one single bivariate copula anymore. Rather, we rewrite $f_{1,2,3}$ bymeans of conditional probability densities (like in Eq. (S1)), *i.e.*,

$$f_{1,2,3}(w_1, w_2, w_3) = f_{3;1,2}(w_3; w_1, w_2) f_{2;1}(w_2; w_1) f_1(w_1), \quad \text{for all } w_1, w_2, w_3 \in \mathbb{R}. \tag{S11}$$

Using Eqs. (S2) and (S6), we get that the second factor $f_{2;1}$ in Eq. (S11) is given by

$$f_{2;1}(w_1; w_2) = \frac{f_{1,2}(w_1, w_2)}{f_1(w_1)} = c_{1,2}\big(F_1(w_1), F_2(w_2)\big)\ f_2(w_2), \quad \text{for all } w_1, w_2 \in \mathbb{R}. \tag{S12}$$

Furthermore, using Eqs. (S5) and (S8), the first factor $f_{3;1,2}$ in Eq. (S11) can be written as follows:

$$\begin{aligned}
f_{3;1,2}(w_3; w_1, w_2) &= \frac{f_{1,3;2}(w_1, w_3; w_2)}{f_{1;2}(w_1; w_2)} \\
&= \frac{c_{1,3;2}\big(F_{1;2}(w_1; w_2), F_{3;2}(w_3; w_2)\big)\ f_{1;2}(w_1; w_2)\ f_{3;2}(w_3; w_2)}{f_{1;2}(w_1; w_2)} \\
&= c_{1,3;2}\big(F_{1;2}(w_1; w_2), F_{3;2}(w_3; w_2)\big)\ f_{3;2}(w_3; w_2),
\end{aligned}$$

for all $w_1, w_2, w_3 \in \mathbb{R}$. Therein, for the factor $f_{3;2}$ we get that

$$f_{3;2}(w_3; w_2) = \frac{f_{2,3}(w_2, w_3)}{f_2(w_2)} = c_{2,3}(F_2(w_3), F_3(w_3))\ f_3(w_3), \quad \text{for all } w_1, w_2, w_3 \in \mathbb{R},$$

by Eq. (S6). Inserting these expressions for $f_{2;1}$ and $f_{3;1,2}$ into Eq. (S11) yields

$$\begin{aligned}
f_{1,2,3}(w_1, w_2, w_3) &= f_1(w_1)\ c_{1,2}\big(F_1(w_1), F_2(w_2)\big)\ f_2(w_2) \\
&\quad \times c_{1,3;2}\big(F_{1;3}(w_1; w_3), F_{3;2}(w_3; w_2)\big)\ f_{3;2}(w_3; w_2) \\
&= f_1(w_1)\ c_{1,2}\big(F_1(w_1), F_2(w_2)\big)\ f_2(w_2) \\
&\quad \times c_{1,3;2}\big(F_{1;2}(w_1; w_2), F_{3;2}(w_3; w_2)\big)\ c_{2,3}\big(F_2(w_2), F_3(w_3)\big)\ f_3(w_3),
\end{aligned}$$

for all $w_1, w_2, x_3 \in \mathbb{R}$. Thus, in a more compact form that suppresses the arguments of the functions, we have

$$f_{1,2,3} = c_{1,2}(F_1, F_2)\ c_{2,3}(F_2, F_3)\ c_{1,3;2}(F_{1;2}, F_{3;2}) \prod_{k=1}^{3} f_k. \tag{S13}$$

Finally, the fourth variable W_4 is added, where we start again with rewriting $f_{1,2,3,4}$ by means of univariate conditional densities, *i.e.*,

$$f_{1,2,3,4}(w_1, w_2, w_3, w_4) = f_{4;1,2,3}(w_4; w_1, w_2, w_3)\ f_{3;2,1}(w_3; w_1, w_2)\ f_{2;1}(w_2; w_1)\ f_1(w_1), \tag{S14}$$

for all $w_1, \ldots, w_4 \in \mathbb{R}$. Comparing this expression with Eq. (S11) we can see that we have supplied already all factors except $f_{4;1,2,3}$. Using Eqs. (S5) and (S8), this conditional probability density takes the form

$$
\begin{aligned}
f_{4;1,2,3}(w_4; w_1, w_2, w_3) &= \frac{f_{1,4;2,3}(w_1, w_4; w_2, w_3)}{f_{1;2,3}(w_1; w_2, w_3)} \\
&= \frac{c_{1,4;2,3}\big(F_{1;2,3}(w_1; w_2, w_3), F_{4;2,3}(w_4; w_2, w_3)\big)}{f_{1;2,3}(w_1; w_2, w_3)} \\
&\quad \times f_{1;2,3}(w_1; w_2, w_3) f_{4;2,3}(w_4; w_2, w_3) \\
&= c_{1,4;2,3}\big(F_{1;2,3}(w_1; w_2, w_3), F_{4;2,3}(w_4; w_2, w_3)\big)\ f_{4;2,3}(w_4; w_2, w_3),
\end{aligned}
$$

for all $w_1, \ldots, w_4 \in \mathbb{R}$, where

$$
\begin{aligned}
f_{4;2,3}(w_4; w_2, w_3) &= \frac{f_{2,4;3}(w_2, w_3; w_4)}{f_{2;3}(w_2; w_3)} \\
&= \frac{c_{2,4;3}(F_{2;3}(w_2; w_3), F_{4;3}(w_4; w_3))\ f_{2;3}(w_2; w_3)\ f_{4;3}(w_4; w_3)}{f_{2;3}(w_2; w_3)} \\
&= c_{2,4;3}(F_{2;3}(w_2; w_3), F_{4;3})(w_4; w_3)\ f_{4;3}(w_4; w_3),
\end{aligned}
$$

and

$$
\begin{aligned}
f_{4;3}(w_4; w_3) &= \frac{f_{3,4}(w_3, w_4)}{f_3(w_3)} \\
&= \frac{c_{3,4}\big(F_3(w_3), F_4(w_4)\big)\ f_3(w_3)\ f_4(w_4)}{f_3(w_3)} \\
&= c_{3,4}(F_3(w_3), F_4(w_4))\ f_4(w_4),
\end{aligned}
$$

for all $w_2, w_3, w_4 \in \mathbb{R}$. Inserting all these expressions into Eq. (S14) and sorting the factors gives the final decomposition of $f_{1,2,3,4}$ into a product of univariate (marginal) probability densities and bivariate copula densities:

$$
\begin{aligned}
f_{1,2,3,4} &= c_{1,2}(F_1, F_2)\ c_{2,3}(F_2, F_3)\ c_{3,4}(F_3, F_4)\ c_{1,3;2}(F_{1;2}, F_{3;2}) \\
&\quad \times c_{2,4;3}(F_{2;3}, F_{4;3})\ c_{1,4;2,3}(F_{1;2,3}, F_{4;2,3}) \prod_{k=1}^{4} f_k.
\end{aligned} \tag{S15}
$$

Now the multivariate density $f_{1,2,3,4}$ is expressed by bivariate copula densities and the univariate marginal densities. Arranging all factors of this decomposition in Figure S1 shows that we again generated a tree-like representation of related factors. In the spirit of Figure 4in the main text, colored patches highlight the factors, which join the decomposition with each additional variable.

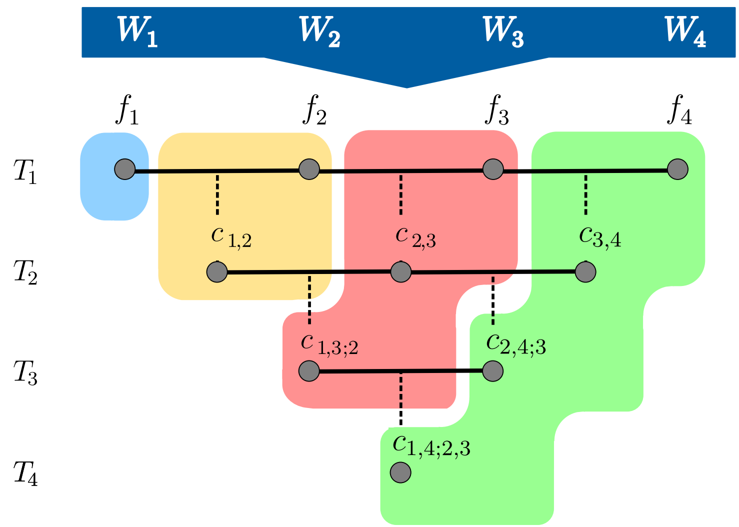

FIGURE S1. Factors contributing to the pair copula decomposition of the quadrivariate probability density $f_{1,2,3,4}$ in Eq. (S15), arranged in a tree representation. Each colored patch contains the factors that are introduced with a new variable.

S2. Modeling of univariate probability densities. Detailed information on the univariate probability densities, which are used in this study to model the (marginal) distributions of microstructure descriptors $\varepsilon, \delta, S_V, \tau_0$ and τ_3, is given in Table S1. Here, I_0 denotes the zero-order modified Bessel function of the first kind, Γ is the Gamma-function, and $\mathbb{1}_A$ denotes the indicator function of a set A. Note that the Rician distribution is the distribution of the length of a two-dimensional random vector with independent and identically (normally) distributed components, see *e.g.* [5].

TABLE S1. Probability density functions used to model the univariate distributions of the microstructure descriptors $\varepsilon, \delta, S_V, \tau_0, \tau_3$, see also [6].

distribution	density function	descriptor	estimated parameters
beta	$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}\mathbb{1}_{[0,1]}(x)$	ε (uncompressed)	$\alpha = 165.297$ $\beta = 244.86$
		ε (compressed)	$\alpha = 109.653$ $b = 227.68$
shifted Gamma	$\frac{(x-1)^{k-1}}{\theta^k\Gamma(k)} e^{-\left(\frac{x-1}{\theta}\right)^k}\mathbb{1}_{[0,\infty)}(x)$	τ_0 (uncompressed)	$k = 43.860$ $\theta = 0.015$
		τ_0 (compressed)	$k = 55.562$ $\theta = 0.011$
		τ_3 (compressed)	$k = 10.740$ $\theta = 0.194$
		τ_3 (uncompressed)	$k = 3.213$ $\theta = 1.537$
Gamma	$\frac{x^{k-1}}{\theta^k\Gamma(k)} e^{-\left(\frac{x}{\theta}\right)^k}\mathbb{1}_{[0,\infty)}(x)$	δ (uncompressed)	$k = 62.9873$ $\theta = 1.6961$
Weibull	$\frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1}\exp\left(-\left(\frac{x}{\alpha}\right)^{\beta}\right)\mathbb{1}_{[0,\infty)}(x)$	δ (compressed)	$\alpha = 98.508$ $\beta = 17.631$
		S_V (compressed)	$\alpha = 0.166$ $\beta = 42.941$
Rician	$\frac{x}{\sigma^2} I_0\left(\frac{xs}{\sigma^2}\right)\exp\left(\frac{-(x^2+s^2)}{2\sigma^2}\right)\mathbb{1}_{[0,\infty)}(x)$	S_V (uncompressed)	$s = 0.162$ $\sigma = 0.004$

S3. Prediction of the Bruggeman exponent. Besides the results presented in Section 3.3, we provide additional information on the comparison between the linear models given in Eqs. [16] and [18], respectively, for the predicting the Bruggeman exponent a. Figure S2 shows scatter plots of the respective residuals, *i.e.*, the difference between the value of a calculated through Eq. [15] and the value of a predicted through the linear models in Eqs. [16] and [18], respectively, in dependence of the specific surface area S_V.

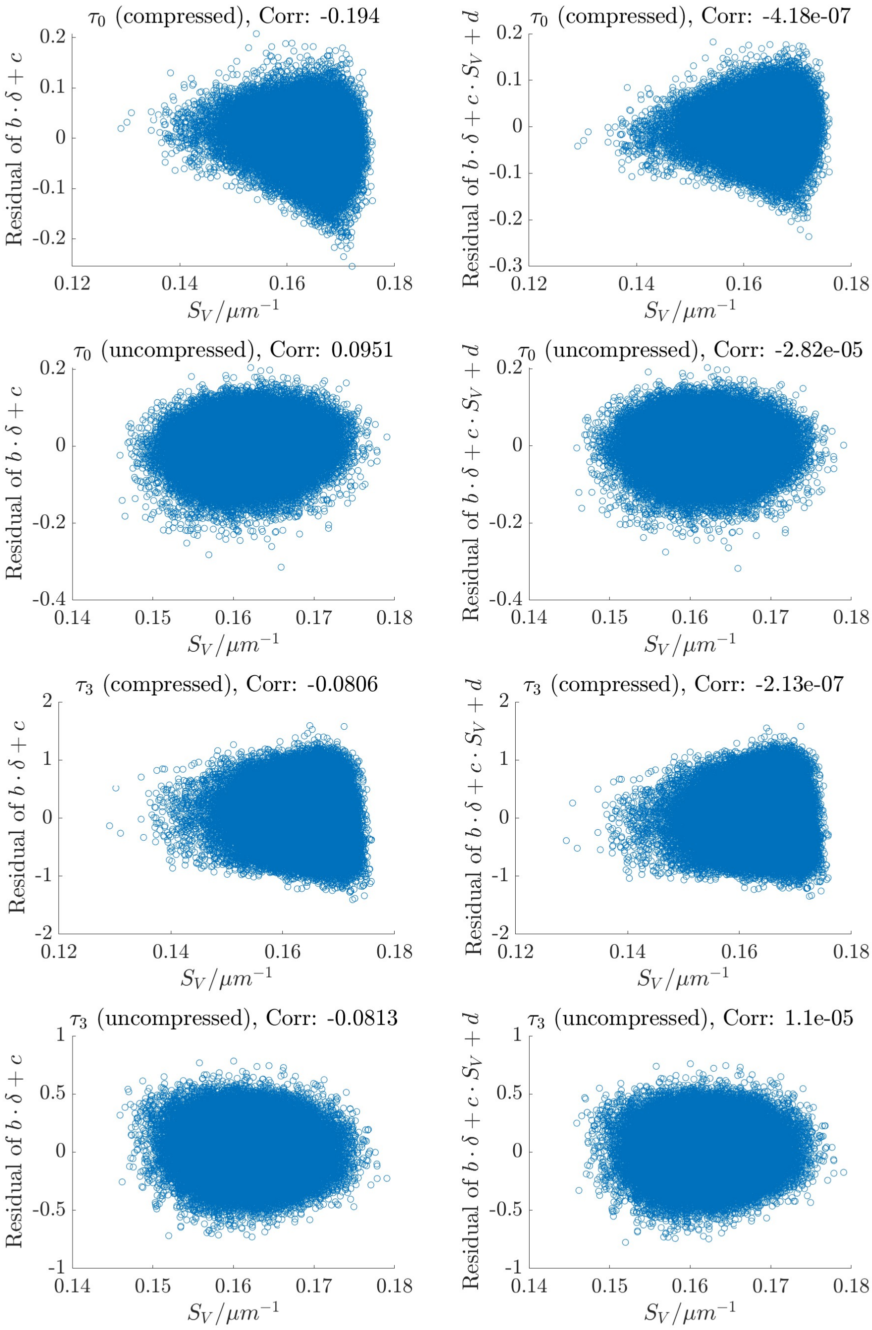

FIGURE S2. Scatter plots of residuals when predicting the Bruggeman exponent a by means of the sheet thickness δ (left column), and by means of δ and the specific surface area S_V (right column). In the line above the plots, the corresponding Pearson correlation coefficient between the residuals and the specific surface area S_V is given.

S4. Bivariate probability densities for descriptor pairs. Besides the results presented in Section 2.5, we provide the bivariate probability densities for all pairs of microstructure descriptors obtained via kernel density estimation based on (i) tomographic image data and (ii) simulated data drawn from the copula-based model, see Figures S3 and S4. The chosen copula families and their associated copula parameters are given in Tables S2 and S3 for the uncompressed and the compressed sample, respectively.

TABLE S2. Copula families and their associated copula parameters in the case of uncompressed paper sheets.

Copula	Model type	Parameters
C_{p,τ_3}	Student-t	$\theta_1 = -0.37, \theta_2 = 30$
C_{τ_0,τ_3}	Survival Gumbel	$\theta = 1.87$
C_{δ,τ_0}	Gaussian	$\theta = -0.41$
$C_{S,\delta}$	Gaussian	$\theta = -0.37$
$C_{p,\tau_0;\tau_3}$	Frank	$\theta = -1.09$
$C_{\delta,\tau_3;\tau_0}$	Independence	
$C_{S,\tau_0;\delta}$	Independence	
$C_{p,\delta;\tau_0,\tau_3}$	Clayton (rotated 270°)	$\theta = -0.27$
$C_{S,\tau_3;\delta,\tau_0}$	Survival Gumbel	$\theta = 1.12$
$C_{p,S;\delta,\tau_0,\tau_3}$	Tawn Type 1 (rotated 270°)	$\theta_1 = -1.49, \theta_2 = 0.08$

TABLE S3. Copula families and their associated copula parameters in the case of compressed paper sheets.

Copula	Model type	Parameters
C_{p,τ_3}	Tawn Type 1 (rotated 270°)	$\theta_1 = -1.38, \theta_2 = 0.19$
$C_{p,\delta}$	BB8(rotated 90°)	$\theta_1 = -4.63, \theta_2 = -0.66$
$C_{p,S}$	Survival BB8	$\theta_1 = 3.55, \theta_2 = 0.91$
C_{δ,τ_0}	BB8(rotated 90°)	$\theta_1 = -2.13, \theta_2 = 0.91$
$C_{S,\tau_3;p}$	Survival Clayton	$\theta = 0.20$
$C_{\delta,\tau_3;p}$	Joe (rotated 90°)	$\theta = -1.25$
$C_{p,\tau_0;\delta}$	BB7 (rotated 90°)	$\theta_1 = -1.04, \theta_2 = -0.45$
$C_{S,\delta;p,\tau_3}$	Independence	
$C_{\tau_0,\tau_3;p,\delta}$	Clayton	$\theta = 0.24$
$C_{S,\tau_0;p,\delta,\tau_3}$	Tawn Type 2	$\theta_1 = 4.89, \theta_2 = 0.00$

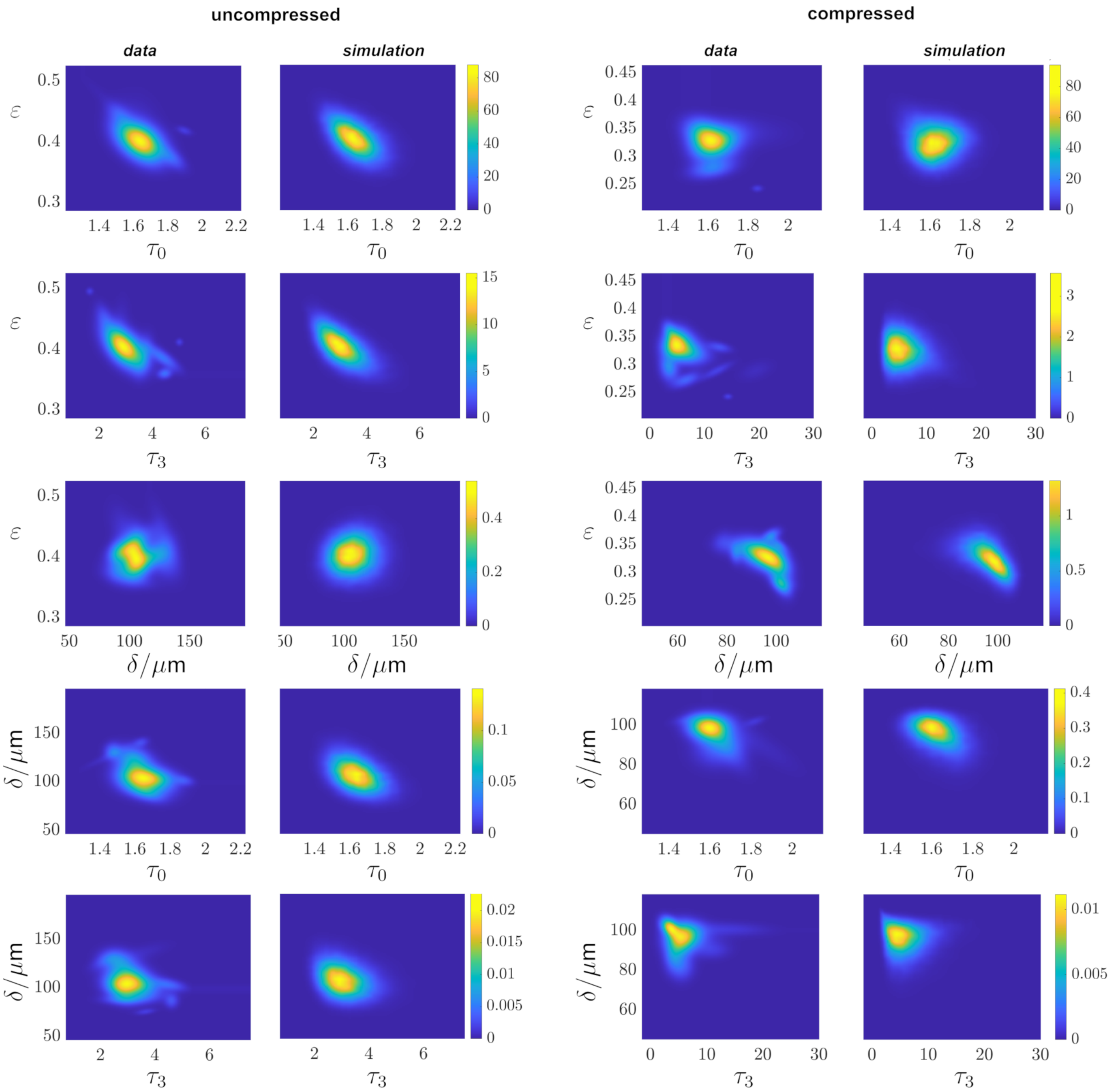

FIGURE S3. Joint probability density of selected descriptor pairs for the uncompressed (left columns) and the compressed sample (right columns). For each descriptor pair, the bivariate density obtained via kernel density estimation is shown for measured image data (left) and simulated data drawn from the copula-based model (right). Shown are descriptor pairs that have a clearly visible negative correlation: (ε, τ_0), (ε, τ_3), (ε, δ), (δ, τ_0), and (δ, τ_3), see also Figure 6 in the main text.

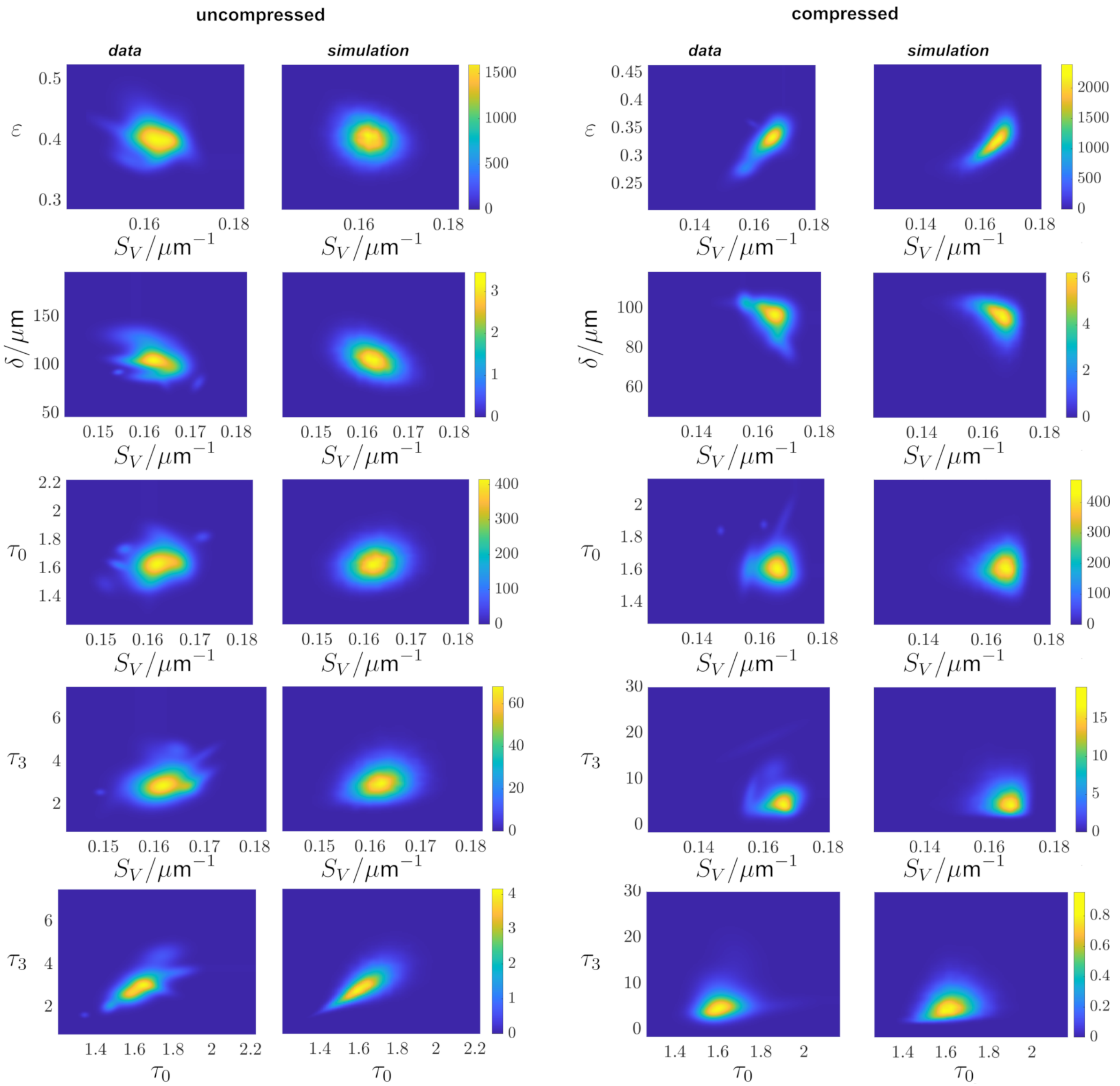

FIGURE S4. Joint probability density of selected descriptor pairs for the uncompressed (left columns) and the compressed sample (right columns). For each descriptor pair, the bivariate density obtained via kernel density estimation is shown for measured image data (left) and simulated data drawn from the copula-based model (right). Shown are remaining descriptor pairs not covered in Figure S3, i.e., (τ_3, τ_0) and all possible pairs containing the specific surface area S_V, see also Figure 6 in the main text.

References

[1] C. Genest. Frank's family of bivariate distributions. *Biometrika*, 74:549–555, 1987.

[2] H. Joe. *Multivariate Models and Dependence Concepts*. Chapman & Hall, 1st edition, 1997.

[3] H. Joe. *Dependence Modeling with Copulas*. CRC Press, 2014.

[4] C. Czado. *Analyzing Dependent Data with Vine Copulas*. Springer, 2019.

[5] Talukdar, K. K. & Lawing, W. D. Estimation of the parameters of the Rice distribution. *The Journal of the Acoustical Society of America*, **89,** 1193–1197 (1991).

[6] Johnson, N. L., Kotz, S.& Balakrishnan, N. *Continuous Univariate Distributions*. J. Wiley & Sons, 1995.