We added the following citations:

- Shrestha2020 -> P. Shrestha and B. Stoeber, Physics of Fluids 32, 011905 (2020).
- Nec2020 -> Y. Nec and G. Huculak, Physics of Fluids 32, 077108 (2020).
- Maggiolo2021 -> D. Maggiolo and S. Sasic, Physics of Fluids 33, 083305 (2021).
- Germanou2020 -> L. Germanou, M. T. Ho, Y. Zhang, and L. Wu, Physics of Fluids 32, 102012 (2020).
- Solano2022 -> T. Solano, C. Ni, R. Mittal, and K. Shoele, Physics of Fluids 34, 051902 (2022).

Reply to Referee #2:

Fisrt, Fig.2 in page 7 was plotted as a function of the relative permeability (kr) while the only liquid phase incopressible three glycerol-water solutions were used in the experiments. You may use the extrapolation [31] and relative error correction [32] methods previously suggested by the authors. Here, I have not clearly understand why you use the relative permeability ratio instead of the absolute permeability. Moreover, how you can convince the accuracy of the corrected relative permeability values. At least, you should include brief introduction about the correction method with its reliability.

Our study deals with single-phase (or absolute) permeability only, including Figure 2. We added the word “absolute” next to “permeability” in several places to emphasize that we operate with absolute permeability in this study. Extrapolation [31] and relative error correction [32] methods were derived from single-phase flow simulations, and they are also applied in this study for single-phase permeability. There is already a detailed discussion of the error correction in Appx. C3.

From the experimental point of view, each mixture of glycerol with water still results in a single-phase solution. We use different ratios of glycerol:water to prepare several single-phase solutions with different viscosities (Fig. 16, Appx, A7), and then we use each solution separately to determine absolute permeability. Absolute permeability value for each solution has to match each other, and this is actually observed in Figure 4.

However, in the text we operate with the term “relative error”, which is calculated for permeability value(s). In some cases, the use of phrase “relative permeability error” may and do lead to confusion. Therefore, we re-checked the manuscript and replaced all “relative permeability error” occurrences with “relative error in permeability” to avoid ambiguity. As mentioned, Fig. 2 presents absolute permeability values, as well the whole study.

Second, the samples used in the study seems to be highly permeable porous media, mostly ranging more than hundreds of Darcy scale, and simple spherical bead shapes. I am curious whether we can generally apply the conclusions from the examinations targetting simple and highly permeable porous media in addition to the correction methods. I would think the key considerations and limits when applying the methods to the real complex and dense porous media should be discussed.

In this study we operate with dimensionless quantities making our analysis universal. All results will hold quantitatively until Stokes physics is valid. Apparent simplicity of the pore space allowed us to obtain accurate match between experiment and simulations, and to clearly justify the use the model system to study the permeability error (Section III, D). We did not use the fact about sphericity of grains to correct our flow simulations. This fact was only helpful to compare our results with other studies (and to show that we are similar in some cases) or to evaluate image contrast. The latter was possible because we knew in advance that there is no bridging between grains, and if there are visible bridges, this is clear imaging artifact.

Additionally, the knowledge about average bead size was helpful to define highest target resolution range (60+ voxels per bead diameter) at the stage of experiment planning (Appx. A1). This was needed to perform extrapolation of the simulated permeabilities. Taking other homogeneous porous media comprised of non-spherical grains but of similar porosity will show identical results. Extrapolation originates not from the grain shape but from the interplay between bulk and boundary error contributions of the numerical scheme, and their limiting behavior as r approaches infinity.

Once the permeability error analysis was performed, and pressure loss at the sample–tube contact was demonstrated (and its possible impact on the measured permeability values), the overall picture is quite universal: one needs to have only external pressure ports connected to the in(eff)fluent tubing of the diameter smaller than the porous sample diameter. Whether the sample is of spherical beads or other grains, or non-granular — it is not important. To demonstrate this point, we took some random sandstone sample and observed the expected discrepancy in permeability between external and internal pressure ports (Fig. 7).

We added the following text:

“In the current study we operate with dimensionless permeability making our results quantitatively universal as long as Stokes physics is valid. We assume that deviations from it, such as occurrence of inertial or Klinkenberg effects, will not change qualitatively the reported results.”

Next, in page 10, the match procedures were explained. However, the explanations how the 35% error has been decreased to 0% in Fig.4 were not systematical and unclear. If the errors were originated from each factor you described in the previous section, you'd better to explain each factor step by step with equations (using tables and flow charts). Particularly, one of the key factors you emphasized was the area ratios (A_tube / A_sample). Here, each region (tube or plug ...) has different length (corrected porosity). So, I am currious which values were used to correct the experimental results. Obviously, the correction procedures and equations should be presented with tables summarizing the results in each step.

Good question. It seems that this point was not explicitly stated. In subsection B of III Results, experimental permeability values are calculated using the pressure gradient (Pin-P0)/Lsystem, contrary to the (Pin-P0)/Lsample used in subsection A of III Results. This is the only difference between

interpretation of the two experimental cases, initial comparison (III Results, A) and the match (III Results, B). We added the following phrase:

" The corresponding experimental permeability values are now obtained using not Lsample but Lsystem: the pressure gradient is calculated as (Pin−P0)/Lsystem. "

Adding tubing and plugs to the simulation setup and extracting superficial velocity from the simulated flow field requires calculation of the porosity of the whole system (sample + plugs + tubing), which is not trivial point. But this is already explained in subsection C of III Results.

In particular, the final results using the correction method were summarized in the end parts (page 16 - 17) of the main text. In line 360, the emperical correction equation is presented as a part of text. It seems not a general way to summarize the key findings in the research manuscript.

We do not consider the presented very approximate and partially working empirical correlation as a main result of our work. Main results are: what steps are necessary to match experiment and simulation from first principles, existence of a virtual geometry of constant cross section, existence of (possibly severe) pressure loss at the tube-sample contact.

Nevertheless, we dedicated a separate line for the presented empirical correlation, but we also added emphasis that this correlation is very approximate.

And the following phrase in the Outlook section was added:

"The error magnitude can be very approximately estimated using k_smpl/k_incorrect − 1 = 0.5(Dsmpl/Dtube)/(Lsmpl/Dsmpl) empirical correlation."

Moreover, I am wodering whether the correction could be general because the combinations of the key two terms can have many numbers of cases. For example, in case that the (D_s / D_t) term is accidentally the exact reciprocal of the latter term (L_s / D_s), the E results in 0.5 always. Herein, I am still curiou if many different combinations of D_t and L_s, which also set E as 0.5, are always satisfy the correct rule suggested in this study.

As we mentioned, the presented empirical correlation is -very- approximate, and it already fails when the tubing diameter equals sample diameter (Dsmpl/Dtube = 1) because there is no pressure loss at tube-sample contact. But for the studied set of parameters (Dsmpl/Dtube > 2) it gives an idea about the possible flaw magnitude.

Concerning the case of the constant product of (Dsmpl/Dtube)*(Dsmpl/Lsmpl), which is, for example, may be equal to 1, resulting in the error E = 0.5 * 1 = 0.5 — there is nothing wrong with this case, the overestimation will be always +50%. I.e., for some ratio of sample-to-tubing diameters one always can find such sample length that the resulting E will be +50%.

Reply to Referee #3:

The manuscript considers sources of error for estimating permeability, for both experiments and simulations, although mainly focused on the experimental setup and imaging. The manuscript is overall extremely thorough and sheds light on many (often neglected) sources of errors. The results of the study are very useful not only for the experimental community, but also for those dealing with image segmentation of experimental samples. The manuscript contains many technical details, which is nicely organized in appendices, making the main part of the manuscript still manageable to read. Overall I only have very minor comments where the authors should clarify some things.

1. Fig. 2 is a bit confusing since k^{ref}_{p3} is not properly defined, and neither what "optimal" settings are. This does not necessarily have to be explained in the figure caption, but somewhere in the text (potentially by including a reference to the appendix where this information can be found) would be good.

After detailed investigation of this comment we found that the original version of Fig. 2 is confusing: the Y-label is k_P3 / k_P3^ref – 1 while the figure shows also computer-generated case, k_beads1 / k_beads1^ref – 1. We modified the figure caption and the figure2 itself, as well as the accompanying text to explicitly mention k_*^ref and “optimal” CT settings.

After re-checking the manuscript, we decided to emphasize the importance of the reference k-values for computer-generated and CT images placing them in a separate line each (Appendix C3, lines 1107+, 1124+).

---

2. In the legend of Fig. 3, it is referred to "P2", but I assume this should be "P3"?

We appreciate the careful examination of our work. Referee3 is right, there has to be P3 (P2 was in older version of the manuscript), and we modified the legend accordingly.