Supplementary Video caption:

Supplementary Video 1: k-Wave forward simulation of PA waves generated by a vertical blood vessel without microbubbles.

Supplementary Video 2: k-Wave forward simulation of PA waves generated by a vertical blood vessel with 1 microbubble.

Supplementary Video 3: k-Wave forward simulation of PA waves generated by a vertical blood vessel with 5 microbubbles.

Supplementary Video 4: PA B-mode images of blood-microbubble mixture flowing in a tube knot with 532 nm illumination.

Supplementary Video 5: PA B-mode images of pure blood flowing in a tube knot with 532 nm illumination.

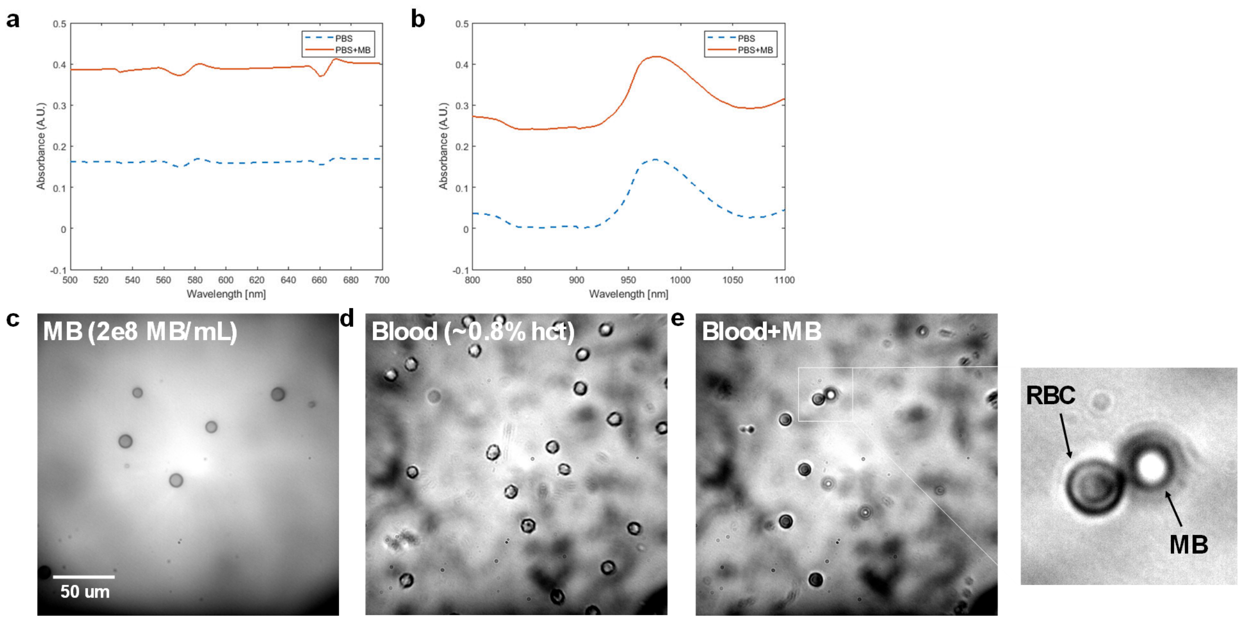
Supplementary Video 6: Doppler US images of the flowing blood-microbubble mixture in a tube knot.

Supplementary Video 7: SVD-filtered PA images of flowing blood-microbubble mixture in a tube knot with 532 nm illumination.

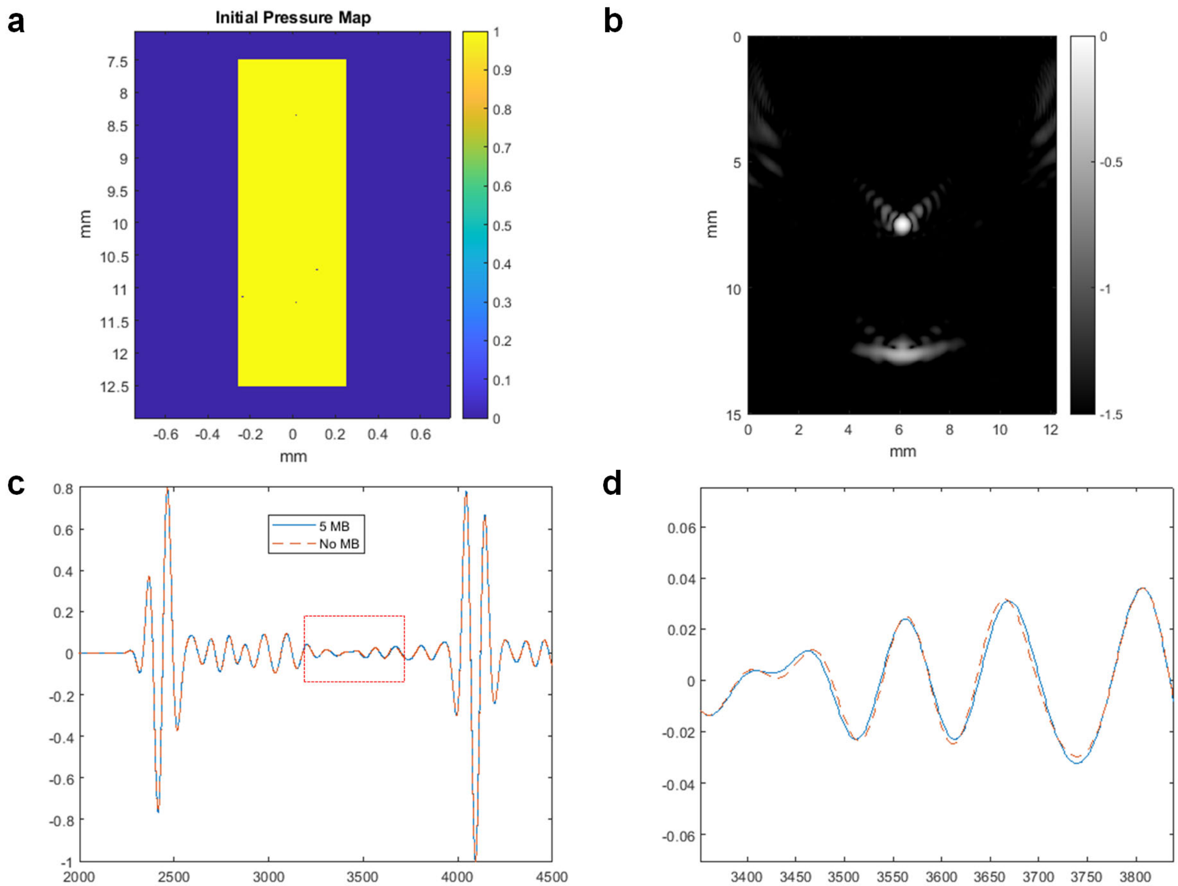
Supplementary Video 8: SVD-filtered PA images of flowing pure blood in a tube knot with 532 nm illumination.

Supplementary Video 9: PA B-mode images of rat liver before and after motion correction.

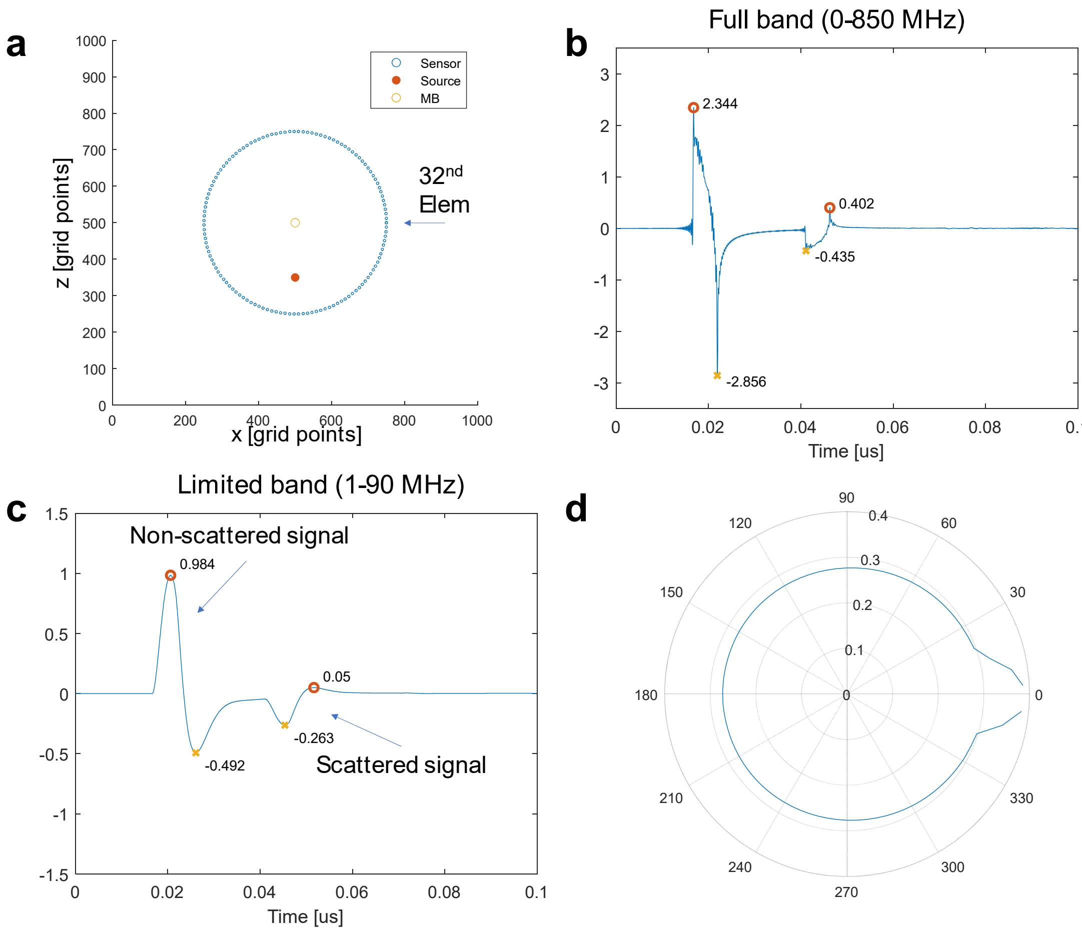
Supplementary Video 10: PA B-mode and SVD-filtered images of rat liver with 3-cycle of hypoxia challenge.



Supplementary Figure 1. Characterization of microbubble properties. (a-b) Microbubble's attenuation coefficient measured using spectrophotometer (Shimadzu, Kyoto, Japan) with transmission mode. The microbubble was prepared by diluting the stock with PBS to form a final concentration of 2×10^7 particles/ml. The PBS+MB has higher attenuation than PBS alone, due to the strong optical scattering by microbubbles. **(c-d)** Microbubble and diluted blood under microscope respectively. **(e)** 2×10^8 particles/mL microbubble was mixed with diluted blood at 0.8% hematocrit and examined under microscope. The microbubble's size is comparable to that of the red blood cell.



Supplementary Figure 2. The impact of microbubble's optical heterogeneity on a vertical vessel visualization. The simulation setup was similar to that in Figure 3, except the acoustic properties of the microbubble were the same as the surrounding blood. By using homogenous acoustic properties within the blood vessel, the simulation results only reflect the optical heterogeneity. **(a)** The vertical blood vessel with randomly distributed microbubbles, and no initial pressure was given to those bubbles due to their lack of optical absorption. **(b)** Reconstructed image with time reversal, assuming homogenous medium. No microbubbles can be visualized on the reconstructed image. **(c-d)** Plot of channel data with 5 and 0 microbubble placed in the vertical vessel. Due to the small size and sparsity of the microbubble, no significant difference is observed for with and without microbubbles. Therefore, comparing the results in Figure 3, we conclude that the acoustic scattering of the microbubbles instead of its lack of optical absorption is the dominating factor for image enhancement of VPS-PAT.



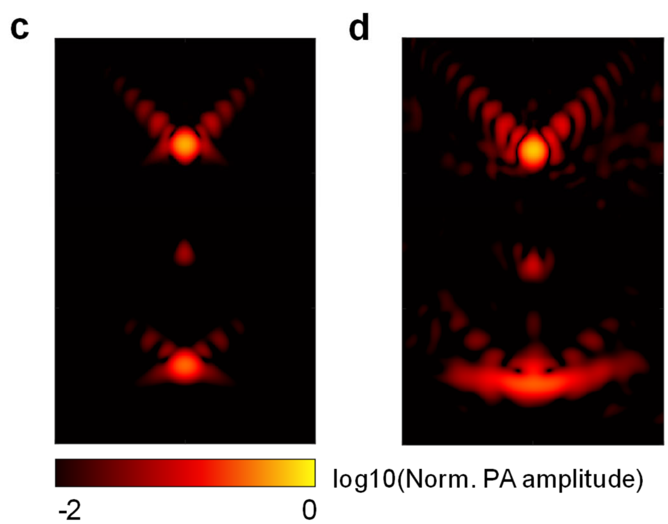
Supplementary Figure 3. Acoustic scattering of the microbubbles. We ran a simplified 2D k-Wave simulation to demonstrate the isotropic scattering of PA signals by microbubbles. **(a)** Illustration of simulation setup. A total of 128 point-detectors were defined around the microbubble. Initial pressure was only provided to the nearby source mimicking the surrounding blood. k-Wave grid size was 0.2 × 0.2 μm^2. **(b-c)** Plot of channel data from the 32^{nd} detector with a full bandwidth and a limited bandwidth. The first signal was the direct signal from the source, and the second signal was the scattered signal by the microbubble. **(d)** Polar distribution of the scattered signal intensity around the microbubble, showing an almost isotropic scattering, i.e, Raman scattering.

a

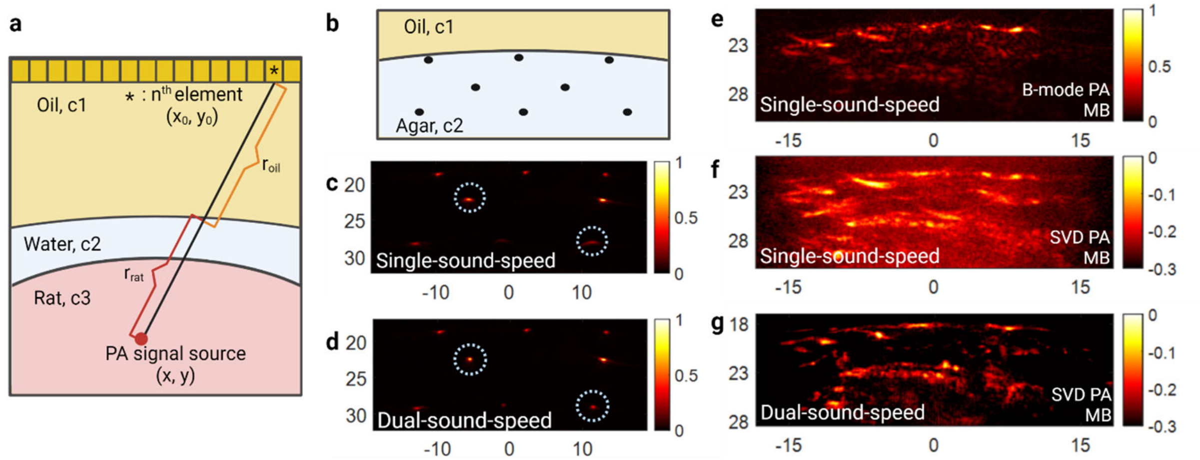
k-wave parameter	Parameter value
Grid size	5 μm
Simulation region size	12.2 by 15 mm^2
Sensor central frequency	5.208 MHz
Sensor bandwidth	4-7 MHz
Sensor spacing	0.298 mm
Microbubble diameter	5 μm
Vessel size	200 μm x 5 mm
CFL	0.3
PML	40

b

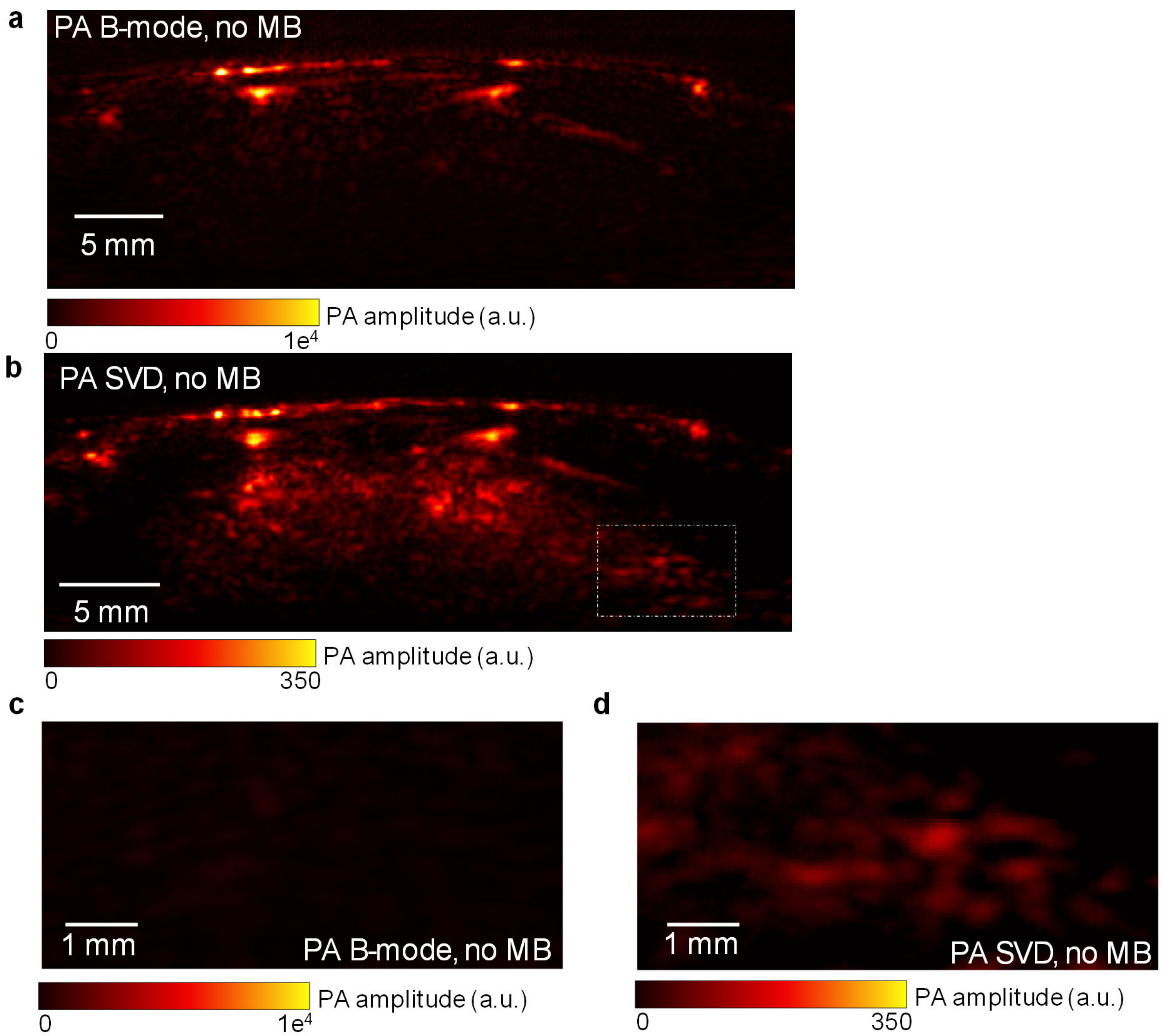
Medium type	Sound speed (m/s)	Density (kg/m^3)	Attenuation (dB/MHz/cm)
Connective tissue	1613	1120	0.5
Fat	1479	937	0.4
Liver	1570	1064	0.5
Muscle	1566	1070	0.15
Tissue	1540	1000	0.5
Blood	1520	1000	0.005
Microbubble	500	10	1



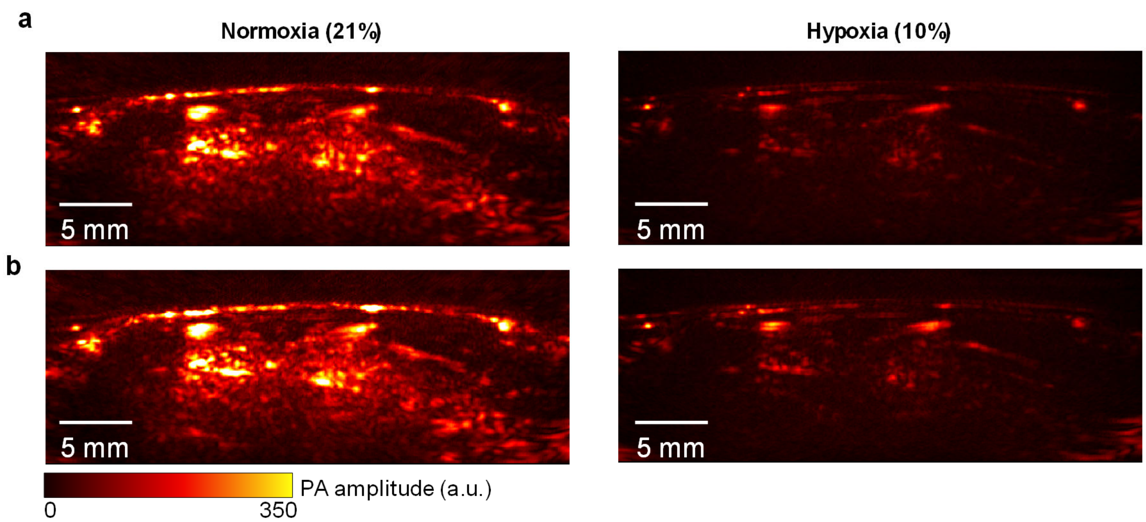
Supplementary Figure 4. Forward k-Wave simulation and reconstruction setup. (a) Defined parameters for k-Wave forward simulation. **(b)** Defined medium acoustic properties for k-Wave forward simulation. **(c-d)** Reconstruction results with the exact sound speed (c) or a homogenous sound speed of 1500 m/s (d).



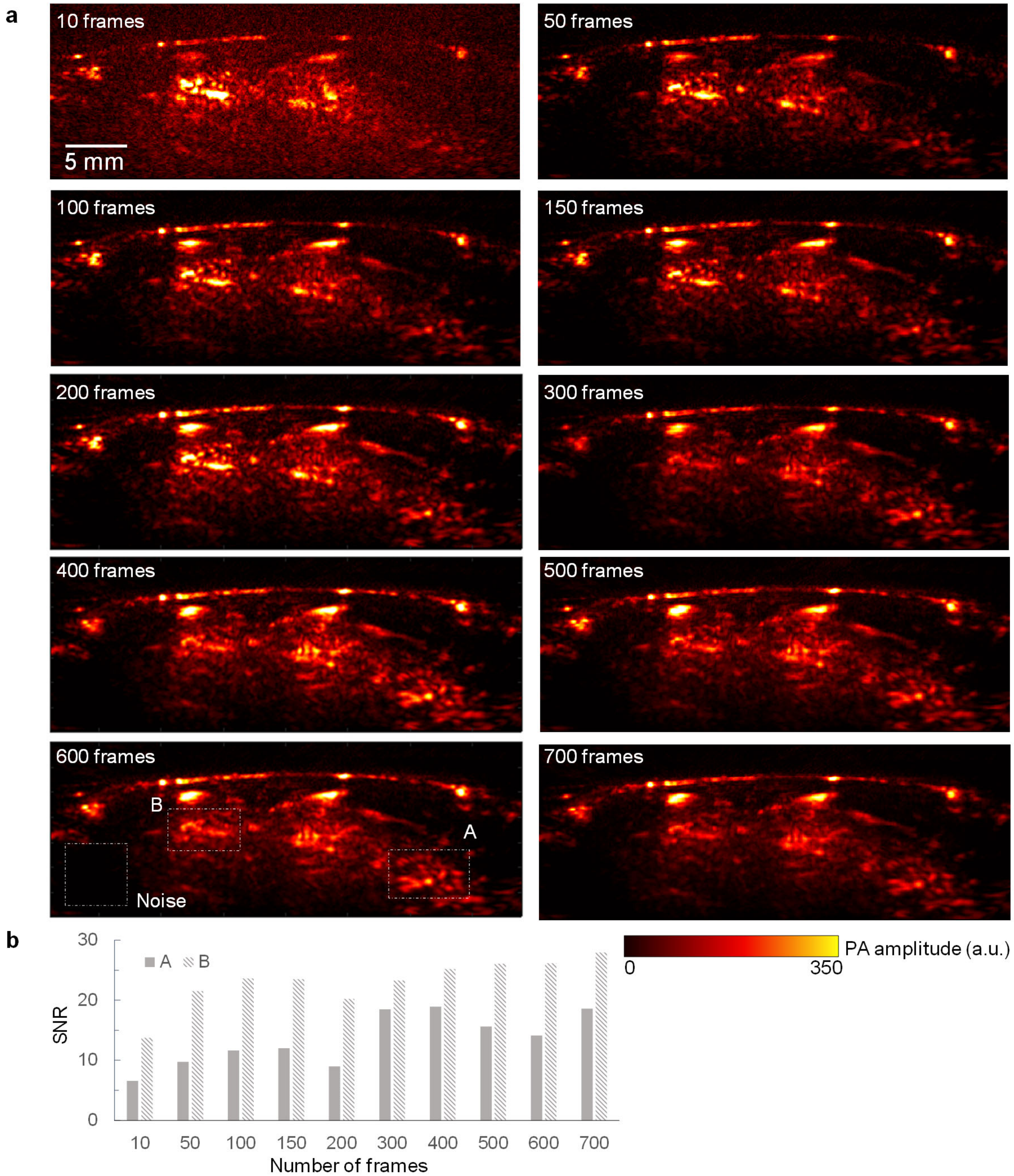
Supplementary Figure 5. Dual-sound-speed (dual-sos) based delay-and-sum (DAS) reconstruction. **(a)** Schematic of *in vivo* imaging system setup for dual-sos reconstruction. The travel distance of PA signals within the tissue was calculated separately for the oil (r_{oil}) region and the water/tissue (r_{rat}) region, with the measured location of the imaging window membrane. **(b)** *In vitro* validation setup for dual-sos DAS. Three layers of single hair strands were placed orthogonal to the imaging plane above the imaging membrane. **(c-d)** B-mode PA images reconstructed by single-sos DAS and dual-sos DAS. The dual-sos DAS provided better signal converging. **(e)** Averaged B-mode PA images by single-sos DAS. **(f-g)** SVD filtered PA images by single-sos DAS and dual-sos DAS.



Supplementary Figure 6. PA B-mode and SVD filtered image of the rat liver without microbubbles. (a-b) B-mode and averaged SVD-filtered PA images of the rat liver region without microbubbles. **(c-d)** Close-up images of the deep region indicated by the dashed box in (b). The SVD filtering had less effect on improving the image quality than the condition with microbubbles.



Supplementary Figure 7: Repeatability of hypoxia challenge. Averaged SVD filtered PA images under normoxia and hypoxia conditions for two additional cycles.



Supplementary Figure 8. Different number of frames used for SVD filtering. SVD filtering were compared using 10, 50, 100, 150, 200, 300, 400, 500, 600, and 700 stabilized frames with similar cutoff, $\lambda_d \approx 0.1\lambda_{max}$. The impact of the number of frames on improving the structure visibility was insignificant beyond 100 frames. Larger number of frames used for SVD can help improve image SNR through stronger averaging, but may also introduce more motion artifact and lead to blurring of small features.