Exploiting flow dynamics for super-resolution in contrast-enhanced ultrasound

Oren Solomon, *Student Member, IEEE,* Ruud J.G. van Sloun, *Member, IEEE*, Hessel Wijkstra, Massimo Mischi *Senior member, IEEE* and Yonina C. Eldar, *Fellow, IEEE*

Abstract—Ultrasound localization microscopy offers new radiation-free diagnostic tools for vascular imaging deep within the tissue. Sequential localization of echoes returned from inert microbubbles with low-concentration within the bloodstream reveal the vasculature with capillary resolution. Despite its high spatial resolution, low microbubble concentrations dictate the acquisition of tens of thousands of images, over the course of several seconds to tens of seconds, to produce a single super-resolved image. Such long acquisition times and stringent constraints on microbubble concentration are undesirable in many clinical scenarios. To address these restrictions, sparsity-based approaches have recently been developed. These methods reduce the total acquisition time dramatically, while maintaining good spatial resolution in settings with considerable microbubble overlap. Here, we further improve sparsity-based super-resolution ultrasound imaging by exploiting the inherent flow of microbubbles and utilize their motion kinematics. While doing so, we also provide quantitative measurements of microbubble velocities. Our method relies on simultaneous tracking and super-localization of individual microbubbles in a frame-by-frame manner, and as such, may be suitable for real-time implementation. We demonstrate the effectiveness of the proposed approach on both simulations and *in-vivo* contrast enhanced human prostate scans, acquired with a clinically approved scanner.

Index Terms—Ultrasound, Contrast agents, Super-resolution, Compressed sensing, Kalman filter.

I. INTRODUCTION

In the past several decades, ultrasonic contrast agents have been utilized successfully in numerous applications [1–3]. In particular, contrast enhanced ultrasound (CEUS) imaging takes advantage of inert microbubbles (MBs) which are injected into the bloodstream, as means to image blood vessels with improved contrast, compared with standard B-mode ultrasound (US) imaging [4]. In recent years, super-resolution US imaging emerged, and enabled the fine visualization and detailed assessment of capillary blood vessels *in vivo* [5–10].

Super-resolution US relies on concepts borrowed from super-resolution fluorescence microscopy techniques such as photo-activated localization microscopy (PALM) and stochastic optical reconstruction microscopy (STORM) [11, 12], which localize individual fluorescing molecules with sub-pixel precision over many frames and sum all localizations to produce a super-resolved image. In CEUS, individual resonating MBs, similar in size to red blood cells, serve as point emitters. Their subsequent localizations are then accumulated to produce the final super-resolved image of the vascular bed with a ten-fold improved spatial resolution compared with standard CEUS imaging. To produce a reliable reconstruction, low MB concentrations are typically used [6, 7], such that in each frame all MBs are well isolated from one another. The localization procedure then amounts to pin-pointing the centroid of a single Gaussian for each detected MB in the captured movie.

Despite yielding a substantial improvement in the spatial resolution, super-resolution ultrasound imaging typically requires tens of thousands of images to produce a single super-resolved image. Acquisition of such a large number of frames results in long scanning durations, leading to poor temporal resolution on the reconstructed sequence. Furthermore, clinical bolus doses injected to human patients result in high overlap between different MBs [13]. These limitations hamper the clinical applicability of localization-based super-resolution techniques.

To overcome the temporal limitation of localization-based super-resolution while not compromising the spatial resolution of the reconstructed image, sparsity-based [14] approaches were recently proposed [15–17]. These approaches favor overlapping MBs to reduce the total acquisition time. As such, sparsity-based methods achieve faster temporal resolution using standard clinical concentrations of MBs. In [15], sparsity-based super-resolution ultrasound hemodynamic imaging (SUSHI), using ultrafast plane-wave acquisition, demonstrated a super-resolved time-lapse movie of 25Hz, showing super-resolved hemodynamic changes in blood flow within a rabbit's kidney. In [16], using a clinically approved scanner with an acquisition rate of 10Hz, a super-resolved image of a human prostate vasculature was demonstrated. In this work, a clear depiction of vascular bifurcations was obtained, although significant MB overlap was present, by performing frame-by-frame sparse localization and subsequent accumulation of all localizations to produce the final super-resolved image.

One major difference between super-resolution in ultrasound and in microscopy is that the point emitters in ul-

trasound are flowing inside the blood vessels, whereas in microscopy fluorescent molecules are fixed to the sub-cellular organelles. Since the motion of individual MBs is not random but rather within blood vessels, this can be exploited to improve the recovery process. Here, we build on our previous results on super-resolution ultrasound imaging [15, 16] by exploiting the flow kinematics of individual MBs as an additional prior in the sparse recovery process.

While previous super-resolution works focused on ultrafast plane-wave image acquisition, e.g. [5, 15, 17], most clinically used scanners are low-rate scanners (10-25Hz). In this work, our aim is to bridge the gap between super-resolution techniques and data obtained from research platforms in laboratory environments, typically low-rate intensity images where significant MB overlap is present. By doing so, as we demonstrate in Section IV, our technique enables practitioners to analyze readily available CEUS scans, and obtain both architectural as well as functional blood flow information. Such analysis can expedite the process of gaining new insights regarding cancer diagnosis [18], treatment, *in-vivo* flow characterization [8] and more.

Our method combines weighted sparse recovery with simultaneous tracking of the individual MBs in the imaging plane. MBs flow inside blood vessels, hence their movement from one frame to the next is structured. Therefore, MBs are more likely to be found in certain areas of the next frame, given their current locations. In our method, each MB track is used to estimate the position of the MBs in the next frame, and since capillary flow is non-turbulent (peak Reynolds number of 0.001) [8, 19], a linear propagation model is used to describe MBs flow from one frame to the next. The accumulated position estimates are then used to form a weighting matrix for weighted sparse recovery which locates the MBs. This allows to favor more likely locations in the sparse recovery process. With the addition of each new frame, the tracks are updated online. We refer to our method as simultaneous sparsity-based super-resolution and tracking, or Triple-SAT. Since our approach tracks individual MBs, it is possible to also estimate their velocities. We provide *in-vivo* super-resolution CEUS imaging of a human prostate, and show that our velocity estimation agrees with previously published results [13]. An illustration of our proposed concept can be seen in Fig. 1.

The methods proposed in this work relate to those presented in [8], in which the authors incorporated an automated detection and tracking mechanism for localized MBs. However, Triple-SAT differs from [8] in the following ways. First, in [8] the automatic tracking algorithm is not used to improve the localization procedure over consecutive frames. Rather, individual MBs were localized over all frames with low MBs concentration, and only then detection and tracking was performed on the localizations to improve the velocities estimation. Here, we use detections from previous frames to improve the detections in the next frame, using sparse recovery to overcome MBs overlap, due to the use of clinical bolus doses. Second, we exploit coarse measurements of MB movements based on optical flow (OF) estimation [20–23] over the captured low-resolution sequence to improve the tracking performance. Thus, we incorporate not only position measure-
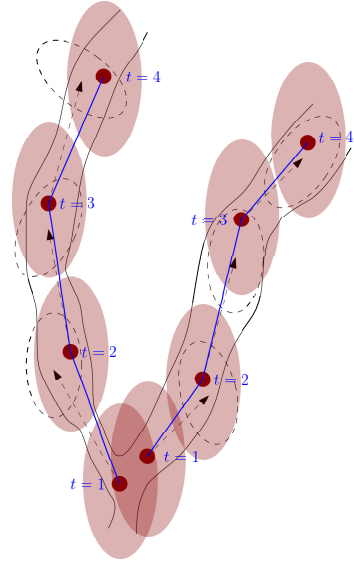

Fig. 1. Proposed concept of Triple-SAT. Individual MBs flow within a blood vessel, depicted here as a bifurcation by black solid contour lines. Large, transparent ellipses represent the echoes measured from individual MBs. In frame $t = 1$ MBs are localized using sparse recovery (small red dots). Using a Kalman filter, their positions are propagated to the next frame ($t = 2$), as marked by the dashed black arrows. Using the error covariance matrix of the filter, ellipses of most likely positions for the MBs are generated, as illustrated by the dashed black ellipses. These ellipses are then used as weights in the sparse recovery process in the next frame ($t = 2$), and so on. The blue lines indicate the estimated trajectories of the MBs.

ments but also velocity measurements in the adopted Kalman filtering framework [24, 25]. These measurements help in improving the overall tracking performance of the MBs, which in turn improves the sparse recovery process. Typically, OF estimation is performed over sequential pairs of images. Here we combine OF estimation with Kalman filtering, as means to include additional information from previous frames and improve the overall estimation performance.

The rest of the paper is organized as follows. In Section II we provide a full description of our method and describe each of its building blocks. Section IV presents *in-silico* as well as *in-vivo* results. In Section V we provide a discussion and conclusions.

Throughput the paper, x represents a scalar, $\mathbf{x}$ a vector, $\mathbf{X}$ a matrix and $\mathbf{I}_{N \times N}$ is the $N \times N$ identity matrix. The notation $||\cdot||_p$ represents the standard p-norm and $||\cdot||_F$ is the Frobenius norm. Subscript x_l denotes the lth element of $\mathbf{x}$ and $\mathbf{x}_l$ is the lth column of $\mathbf{X}$. Superscript $\mathbf{x}^{(p)}$ represents $\mathbf{x}$ at iteration p, T^* denotes the adjoint of T, and $\bar{\mathbf{A}}$ is the complex conjugate of $\mathbf{A}$. The estimated vector in frame k, given the estimate in the $(k-1)$th frame, is indicated by $\mathbf{s}_{k|k-1}$. Likewise $\mathbf{P}_{k|k-1}$ indicates its estimated covariance matrix k, given the $k-1$ estimate. The ijth element of matrix $\mathbf{A}$ is denoted as $A[i, j]$.

II. SIMULTANEOUS SPARSITY-BASED SUPER-RESOLUTION AND TRACKING

A. Principle

In this work, we primarily aim at improving sparsity-based super-resolution ultrasound from movies which were acquired

from low-rate clinical scanners, were we have access only to the final intensity images which are displayed on screen. We start from a contrast-enhanced ultrasound sequence of K frames where each frame consists of $M \times M$ pixels. A contrast-specific imaging mode based on a power-modulation pulse scheme is used to reject tissue signal and enhance signal from MBs [10, 26], such that only MBs are visualized. Prior to Triple-SAT processing, all frames are registered, as described in [16]. After registration, we have K registered low-resolution frames, which serve as the input data to our algorithm.

Figure 2 shows the main flow and building blocks of Triple-SAT. Given the weighting matrix, based on trajectories estimated from the $(k-1)$th frame, we perform weighted sparse recovery to estimate the positions of the MBs on a high-resolution grid in the kth frame. Next, We acquire in the kth frame a crude velocity measurement of the MBs by applying OF estimation on the captured low-resolution sequence. Thus, for each MB, we obtain both positions and corresponding velocities measurements, which are then used in the automatic tracking algorithm to update the positions and velocities of the individually localized MBs using Kalman filtering. The newly estimated positions and velocities are used to generate an updated weighting matrix for sparse recovery of MB positions in the $(k+1)$th frame, while providing quantitative information on the flow kinematics.

The reconstruction process of Triple-SAT can be considered as sparse recovery with time varying support, where the support represents the MBs locations. Previous works on sparse recovery with varying support have been proposed in the compressed sensing literature, such as [27–30]. Triple-SAT differs from previous works in the following manner. First, previous works assume a first-order recursion for the propagation model of the non-zero entries of the sparse signals, i.e. $x_{k+1} = \alpha x_k + v_k$, where x_k is a scalar entry from the sparse vector, v_k is additive Gaussian noise and α is a known constant. In this case, only the support of the sparse signal is of interest, but in CEUS, MBs kinematics also include varying velocities. Here we consider an extended model which does not include only the position estimation of the MBs, but also their velocities. Second, as MBs flow over time, new MBs emerge and some MBs vanish from the imaging plane, due to the 3D geometry of the blood vessels. It is thus desirable to associate new MBs to previous localizations to improve the overall tracking and to achieve a more reliable estimation of their motion kinematics. This association process is not considered in prior works, but is taken into account in Triple-SAT by the use of an automatic association algorithm, combined with Kalman filtering.

We next detail the main building blocks of Triple-SAT.

B. Weighted sparse recovery

We first start with a description of our sparse recovery algorithm. This recovery procedure is performed for each frame separately. Similar to [15], we model each frame as a summation of L_k individual MB echoes,

$$Z_k(x,z) = \sum_{i=1}^{L_k} u(x - x_i, z - z_i)\sigma_i, \tag{1}$$

were $u(\cdot, \cdot)$ is the point spread function (PSF) of the transducer, σ_i is the magnitude of the returned echo from the ith MB located at position x_i, z_i and we assume that the PSF of the transducer is known. This PSF can be measured from the acquired data, as described in [15, 16].

Following similar derivations to [15], we discretize the kth frame in (1) as $\mathbf{Z}_k, k = 1, \ldots, K$ of size $M \times M$, and denote its vectorized form $\mathbf{z}_k$. We also introduce a high-resolution grid of size $N \times N$ pixels, such that $N = PM$ for some $P \geq 1$. We denote the (vectorized) super-resolved frame k, which contains the locations of the MBs on the high-resolution grid, by $\mathbf{i}_k$. Using knowledge of the PSF, the measured frame $\mathbf{z}_k$ is related to the super-resolved frame $\mathbf{i}_k$ via

$$\mathbf{z}_k = \mathbf{H}\mathbf{i}_k, \tag{2}$$

where $\mathbf{H}$ is a known dictionary matrix based on the PSF. We follow [15, 31] and consider recovering $\mathbf{i}_k$ in the discrete Fourier domain. In this domain, $\mathbf{H}$ has the following structure

$$\mathbf{H} = \mathbf{U}(\mathbf{F}_M \otimes \mathbf{F}_M),$$

were $\mathbf{U}$ is an $M \times M$ diagonal matrix, whose diagonal contains the discrete Fourier transformed PSF, and $\mathbf{F}_M$ is a partial $M \times N$ discrete Fourier transform (DFT) matrix, whose M rows contain the M lowest frequencies of a full $N \times N$ DFT matrix. Considering (2) in the discrete Fourier domain leads to a numerically efficient sparse solver, as described in [31]. Estimation of $\mathbf{i}_k$ is achieved by solving the following convex optimization problem,

$$\min_{\mathbf{i}_k \geq 0} ||\mathbf{z}_k - \mathbf{H}\mathbf{i}_k||_2^2 + \lambda||\mathbf{i}_k||_1, \tag{3}$$

where $\lambda \geq 0$ is a regularization parameter.

As in [16], the super-resolved image can be constructed by solving (3) for each frame k and accumulating all localizations. To improve the sparse recovery process, we propose solving the following weighted l_1 minimization problem,

$$\min_{\mathbf{i} \geq \mathbf{0}_k} ||\mathbf{z}_k - \mathbf{H}\mathbf{i}_k||_2^2 + \lambda||\mathbf{W}_k\mathbf{i}_k||_1. \tag{4}$$

The matrix $\mathbf{W}_k$ is an $N^2 \times N^2$ diagonal weights matrix which incorporates the flow dynamics of the MBs in the sparse recovery process, and changes with each frame. Intuitively, this matrix assigns higher weights to locations less probable to contain MBs, thus forces the sparse recovery process to favor specific locations in the frame, which are more likely to contain the MBs. In practice, we minimize (4) using the FISTA algorithm [32, 33], as we describe in Algorithm 1 or by using the reweighted iterative l_1 algorithm [34]. We apply Algorithm 1 for each frame in the movie separately.

In Algorithm 1, L_f is the Lipschitz constant of the quadratic term of (4), readily given as the maximum eigenvalue of $\mathbf{H}^T\mathbf{H}$ and $\mathcal{T}_\alpha$ is the soft-thresholding operator, defined as

$$\mathcal{T}_{\boldsymbol{\alpha}}(\mathbf{x})[i] = \max(0, |x_i| - \alpha_i) \cdot \text{sign}(x_i),$$

where $\boldsymbol{\alpha}$ and $\mathbf{x}$ are vectors of the same length.

We next describe how to construct $\mathbf{W}_k$ per frame using MBs trajectories, Kalman filtering and OF.
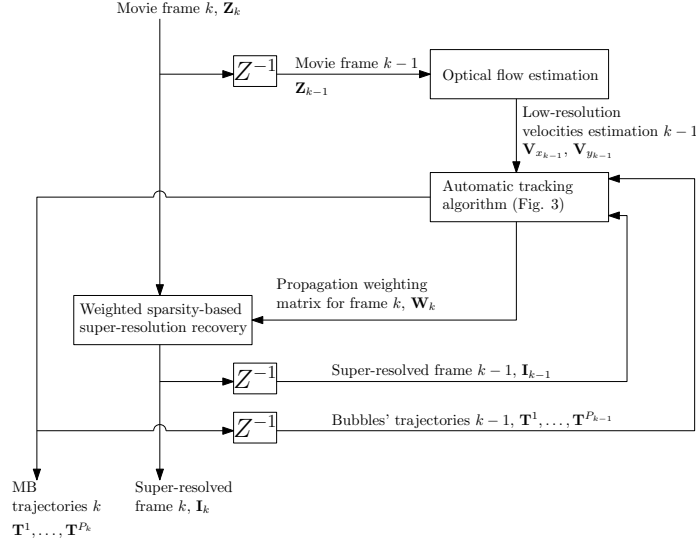
Fig. 2. Main building blocks of Triple-SAT. First, MBs velocities are estimated from frame k using optical flow estimation. The localized MBs from the previous frame $k-1$ are then propagated to frame k, assuming a constant acceleration model using the Kalman filter and the measured velocities obtained from the OF estimation. This procedure yields a new estimate for the true MB positions and velocities, while also forms the weighting matrix for frame k. This weighting matrix is then used in the sparse recovery process to yield the super-resolved frame k. This process repeats itself for each new frame in the movie. Z^{-1} represents a delay of one time-unit.

Algorithm 1 FISTA for minimizing (4)

Require: $\mathbf{z}_k$, $\mathbf{W}_k$, $\lambda > 0$, maximum iterations $D_{\max}$

    Initialize $\mathbf{y}_1 = \mathbf{x}_0 = \mathbf{0}$, $\mathbf{w} = \text{diag}\{\mathbf{W}_k\}$, $t_1 = 1$ and $d = 1$

    while $d \leq D_{\max}$ or stopping criteria not fulfilled **do**

        1: $\mathbf{g}_d = \mathbf{H}^H \mathbf{H} \mathbf{y}_d - \mathbf{H}^H \mathbf{z}_k$

        2: $\mathbf{x}_d = \mathcal{T}_{\frac{\lambda}{L}\mathbf{w}}(\mathbf{y}_d - \frac{1}{L}\mathbf{g}_d)$

        3: Project to the non-negative orthant $\mathbf{x}_d(\mathbf{x}_d < 0) = \mathbf{0}$

        4: $t_{d+1} = \frac{1}{2}(1 + \sqrt{1 + 4t_d^2})$

        5: $\mathbf{y}_{d+1} = \mathbf{x}_d + \frac{t_d - 1}{t_{d+1}}(\mathbf{x}_d - \mathbf{x}_{d-1})$

        6: $d \leftarrow d + 1$

    end while

    return $\mathbf{i}_k = \mathbf{x}_{D_{\max}}$

C. Microbubble tracking

The weighting matrix $\mathbf{W}_k$ represents the accumulated probability of localized MBs from the $(k-1)$th frame to be found in new locations in the kth frame. Its construction requires identifying and tracking individual MBs, as we show in Section II-D. We now turn to explain this process. First, we define the state of the pth MB in frame k, as $\mathbf{s}_k^p \in \mathbb{R}^4$ with

$$\mathbf{s}_k^p = [x_k^p, v_{x_k}^p, y_k^p, v_{y_k}^p]^T.$$

Here, x_k^p and y_k^p are Cartesian coordinates which indicate the position of the pth MB in frame k, and $v_{x_k}^p$ and $v_{y_k}^p$ its respective velocities. The accumulation of all states of the pth MB from frame 1 to frame K_p, $\mathbf{T}^p = [\mathbf{s}_1^p, \ldots, \mathbf{s}_{K_p}^p] \in \mathbb{R}^{4 \times K_p}$, is referred to as the track of the pth MB.

To proceed, we consider an arbitrary frame, k. At this stage, we posses all the states of P previously tracked MBs, $\mathbf{s}_{k-1}^1, \ldots, \mathbf{s}_{k-1}^P$. Given the next low-resolution frame $\mathbf{z}_k$, we have two goals:

1) Recover the locations of the L MBs which are embodied in frame $\mathbf{z}_k$. We note that the number L of MBs in

frame k is generally different than the number of MBs in the previous frame P. This possible discrepancy occurs since blood vessels have a three-dimensional topology, and consequentially MBs may shift in and out of the imaging plane.

2) Associate each newly localized MB to a previously known track, or open a new track if no such association is possible. This association enables to produce the weighting matrix $\mathbf{W}_k$ by propagating the tracks of individual MBs, while it also provides estimation of MB velocities.

The tracking and association process is illustrated in Fig. 3. The output of Algorithm 1 is the $(k-1)$th super-resolved frame, $\mathbf{i}_{k-1}$, whose non-zero values correspond to the positions of the MBs present in this frame. Next, given all previously known tracks $\mathbf{T}^1, \ldots, \mathbf{T}^P$, these positions need to be associated to the tracks. The updated tracks are essential to the formulation of $\mathbf{W}_k$. The goal of the uppermost block in Fig. 3 is to associate each individually localized MB to one of the known P tracks, or to open a new track if no such correspondence is found. Specifically, this matching and association process is realized using the multiple hypothesis tracking (MHT) procedure.

The MHT algorithm, as first suggested by Reid [25, 35], is considered as one of the most popular data association algorithms. The key idea in MHT is to produce a tree of potential hypotheses for each target, in our case MB locations. Upon receiving new measurements, the likelihood of each possible track is calculated and the most likely tracks are selected. This can be done by formulating and solving the maximum weighted independent set [36, 37], for example. The likelihood calculation relies on all past observations of each target [36]. The MHT algorithm is known to produce good data association results due to its pruning stage. Ambiguities

are assumed to be resolvable when new data is acquired. As such, given the latest measurements in frame k, the method estimates the likelihood based upon J previous measurements (where J can be controlled) to resolve past ambiguities in the $(k-J)$th frame irrevocably, and updates all tracks accordingly for the current frame. Thus, data-to-track association decisions are always based upon previous J frames, in a sliding-window manner. An example of associated track numbers to new localizations is presented in Fig. 4, left panel. In practice, we use the Lisbon implementation, taken from [38, 39], which offers full integration into the MATLAB environment.

At the end of this stage, after the association process is finished, existing tracks have been assigned new measurements (MBs positions and velocities), and new tracks are generated, if new MBs were localized. If an existing track was not updated, then this track is closed and cannot be further updated, indicating that the individual MB of this track is no longer present in the movie. We now need to integrate the measurements to their corresponding tracks, and propagate the updated tracks to the next frame k.

Track update and propagation is performed by applying Kalman filtering to each track, individually. Individual tracks represent the history of each localized MB. This history helps to propagate the MBs to the next frame more accurately, and to obtain better velocity estimation. To this end, we consider the pth track. We assume a linear propagation model for the locations of the individual MBs between consecutive frames. This model is given by

$$\mathbf{s}_k^p = \mathbf{\Phi}\mathbf{s}_{k-1}^p + \boldsymbol{\eta}_k^p, \tag{5}$$

where

$$\mathbf{\Phi} = \begin{bmatrix} 1 & \Delta T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \Delta T \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

with $1/\Delta T$ being the frame-rate of the US machine. Model (5) corresponds to the discretized version of the *continuous white noise acceleration (CWNA) model*, or *second-order kinematic model* [40]. Ideally, a constant velocity model has zero acceleration, or zero second-order derivative. In practice, the CWNA model assumes that the velocity of each MB has slight perturbations, modeled as a zero-mean white noise with power spectral density ρ. In the discrete model (5), this uncertainty is captured by the zero-mean additive Gaussian noise vector $\boldsymbol{\eta}_k^p$, associated with a covariance matrix $E\{\boldsymbol{\eta}_k^p \boldsymbol{\eta}_k^{p^T}\} = \mathbf{Q}_k^p$. Following [40], the CWNA model covariance matrix $\mathbf{Q}_k^p$ is given by

$$\mathbf{Q}_k^p = \begin{bmatrix} 1/3\Delta T^3 & 1/2\Delta T^2 & 0 & 0 \\ 1/2\Delta T^2 & \Delta T & 0 & 0 \\ 0 & 0 & 1/3\Delta T^3 & 1/2\Delta T^2 \\ 0 & 0 & 1/2\Delta T^2 & \Delta T \end{bmatrix}\rho,$$

where ρ is chosen empirically.

The measurement model for the pth MB is given by

$$\mathbf{y}_k^p = \mathbf{s}_{k-1}^p + \boldsymbol{\zeta}_k^p, \tag{6}$$

were $\boldsymbol{\zeta}_k^p$ is zero-mean independent i.i.d. Gaussian noise with covariance matrix $E\{\boldsymbol{\zeta}_k^p \boldsymbol{\zeta}_k^{p^T}\} = \mathbf{R}_k^p$. In practice, we choose
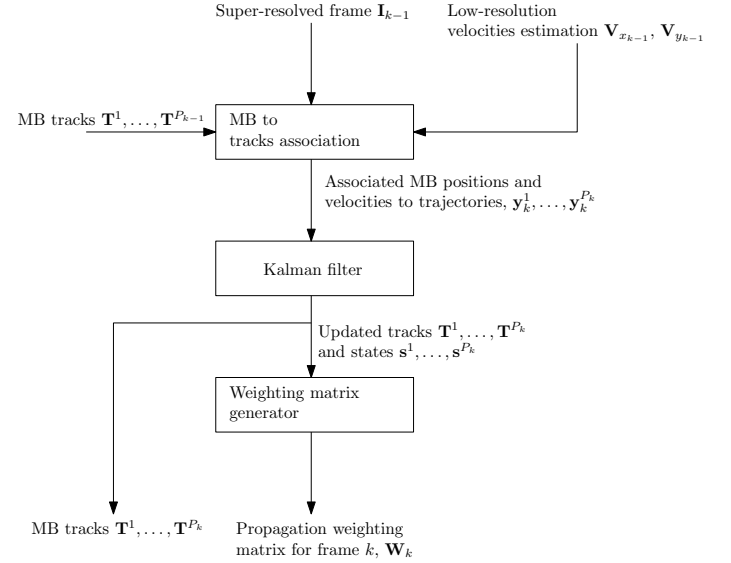

Fig. 3. Automatic tracking and data association procedure. New localized MBs from the $(k-1)$th super-resolved frame are associated to previously known P_{k-1} tracks, or open new tracks, while non-associated tracks are closed (uppermost block, realized by the MHT algorithm). Then, using Kalman filtering, these new P_k tracks are propagated to the next frame (central block). Estimated velocities in the xy plane ($\mathbf{V}_{x_{k-1}}, \mathbf{V}_{y_{k-1}}$) using optical flow on the low resolution $k-1$ frame are associated to the newly localized MBs and used as measured velocities for the Kalman filter update. Thus, an updated track estimation is produced. Lastly, these propagated tracks form the weighting matrix $\mathbf{W}_k$ (lowest block). The tracks serve as inputs to the algorithm in the next frame, when new localizations arrive, and the tracking and association process repeats itself.

this matrix to be a diagonal matrix, and consider its diagonal as free parameters of Triple-SAT, which are chosen empirically.

From the super-resolved image $\mathbf{i}_{k-1}$ we measure the position of the MBs. Specifically, consider an MB which is detected in position $[n_k\Delta_X, n_y\Delta_Y]^T$, where Δ_X and Δ_Y are the known sizes of each pixel in the super-resolved image and $[n_k, n_y]$ are some integers. If the MHT algorithm decided that this MB belongs to the pth track, then $y_k^p[1] = n_k\Delta_X$ and $y_k^p[3] = n_k\Delta_Y$. The velocities of the MBs, or $y_k^p[2]$ and $y_k^p[4]$, are measured using OF estimation [20] on the low-resolution movie frames, as will be described in Section II-E.

Given the propagation (5) and the measurement (6) models, we now formulate the Kalman filter update rules. The model propagation and its corresponding propagated estimation covariance matrix are given by

$$\begin{aligned} \mathbf{s}_{k|k-1}^p &= \mathbf{\Phi}\mathbf{s}_{k-1|k-1}^p, \\ \mathbf{P}_{k|k-1}^p &= \mathbf{\Phi}\mathbf{P}_{k-1|k-1}^p\mathbf{\Phi}^T + \mathbf{Q}_k^p. \end{aligned} \tag{7}$$

Using (7), the weighting matrix $\mathbf{W}_k$ is calculated, as we describe in Section II-D. Next, we solve (4) to recover the kth super-resolved frame, $\mathbf{i}_k$. After the association process is finished, for each track we update its last state via the Kalman filter equations. The Kalman gain is given by

$$\mathbf{K}_k^p = \mathbf{P}_{k|k-1}^p (\mathbf{P}_{k|k-1}^p + \mathbf{R}_k^p)^{-1}, \tag{8}$$

and the innovation step along with the updated estimation error

covariance matrix are given by

$$\begin{aligned}\mathbf{s}^p_{k|k} &= \mathbf{s}^p_{k|k-1} + \mathbf{K}^p_k(\mathbf{y}^p_k - \mathbf{s}^p_{x|x-1}) \\ \mathbf{P}^p_{k|k} &= (\mathbf{I}_{6\times6} - \mathbf{K}^p_k)\mathbf{P}^p_{x|x-1}.\end{aligned} \quad (9)$$

From the innovation step (9), we update the states as $\mathbf{s}^p_k = \mathbf{s}^p_{k|k}$ with estimation covariance matrix $\mathbf{P}^p_k = \mathbf{P}^p_{k|k}$.

D. Weighting matrix formulation

After the states for all MBs are propagated using (7) and associated to existing or new tracks, we turn to formulate the weighting matrix $\mathbf{W}_k$, as illustrated in Fig. 3, lowest panel. The propagated state $\mathbf{s}^p_{k|k-1}$ represents the position and velocity of the pth MB, and has its associated estimation error covariance matrix $\mathbf{P}^p_{k|k-1}$. Based on the state predictions, we formulate a spatial MB-likelihood map $\mathbf{J}_k$, by assigning probabilities drawn from an anisotropic Gaussian distribution of which the mean and covariance are dictated by their respective predictions/updates in the Kalman framework. This process is illustrated in Fig. 5. By aggregating the estimated positions and Gaussians of all of the P propagated MBs, a spatial map of their possible true locations on the high-resolution grid is constructed, denoted as $\mathbf{J}_k$. The ijth element of this $N \times N$ matrix is expressed as

$$J_k[i,j] = \sum_{p=1}^{P} A^p e^{-q^p\left(\frac{1}{\sigma_x^{p2}}(i-x_0^p)^2 - c^p(i-x_0^p)(j-y_0^p) + \frac{1}{\sigma_y^{p2}}(j-y_0^p)^2\right)},$$

with $A^p = \sqrt{|2\pi\mathbf{P}^p_{k|k-1}|}$, $[x_0^p, y_0^p] = [s^p_{k|k-1}[1], s^p_{k|k-1}[3]]$, $\sigma_x^p = P^p_{k|k-1}[1,1]$, $\sigma_y^p = P^p_{k|k-1}[3,3]$, $q^p = 1/(2(1-\rho^{p2}))$, $c^p = 2\rho^p/(\sigma_x^p\sigma_y^p)$ and $\rho^p = P^p_{k|k-1}[1,3]/(\sigma_x^p\sigma_y^p)$. The diagonal of the weighting matrix $\mathbf{W}_k$ is the inverse of the vectorized form of $\mathbf{J}_k$ plus a small regularization value ϵ, to avoid division by zero,

$$\begin{aligned}W_k[i,i] &= \frac{1}{J_k[\lfloor i/N \rfloor, (i \mod N)] + \epsilon}, \\ i &= 1, \dots, N^2,\end{aligned} \quad (10)$$

where $\lfloor \cdot \rfloor$ is the floor operation and $(x \mod y)$ is the modulo operation with the swap $0 \to N$. An example illustration of such a weighting matrix can be observed in Fig. 4, right panel. Vectorization of this $N \times N$ image is the diagonal of $\mathbf{W}_k$. The main building blocks of Triple-SAT are described in Algorithm 2.

As noted before, in (6) we assume that we measure not only the positions of localized MBs, but also their velocities. We now turn to describe how this velocity measurement is done.

E. Velocity regularization via optical flow estimation

To improve the tracking procedure of individual MBs, we provide velocity measurements to the Kalman filter as part of the input to (6). This is done by OF estimation [20–23] from the low-resolution movie. Although the formulation originates from the low-resolution movie, in which individual MBs are not separable, this added velocity information helps in regularizing the tracking of individual MBs. Consider for example, a newly localized MB. This MB has a single position measurement. Without added information of its general
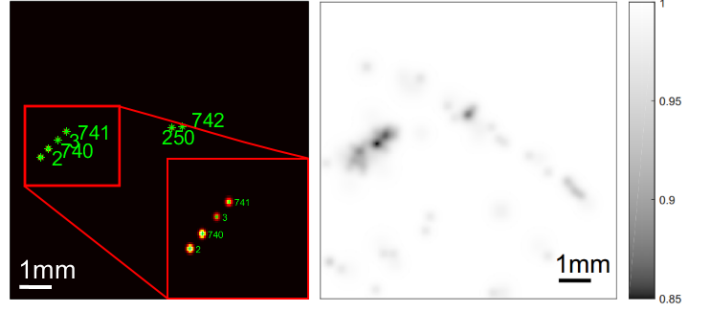

Fig. 4. Left panel shows an example of association of new measurements to existing tracks. Red box indicates enlarged area. MBs were smoothed slightly for visualization purposes only. Right panel depicts an example of the weighting matrix $\mathbf{W}_k$, presented as an $N \times N$ image.
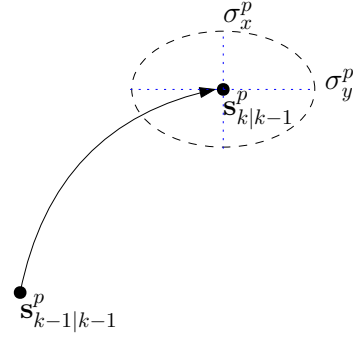

Fig. 5. Generation of the weighting matrix $\mathbf{W}_k$. Previously estimated state $\mathbf{s}^p_{k-1|k-1}$ is propagated to state $\mathbf{s}^p_{k|k-1}$, according to (7). Its propagated error covariance matrix $\mathbf{P}^p_{k|k-1}$ is then used to draw an ellipse around its location, where $\sigma_x^p = \mathbf{P}^p_{k|k-1}[1,1]$ and $\sigma_y^p = \mathbf{P}^p_{k|k-1}[3,3]$. Aggregation of all propagated uncertainty ellipses generates an image of possible MBs locations. $\mathbf{W}_k$ is then proportional to the inverse of this image.

direction of movement, the tracking filter will propagate the MB to a position, which in general is not related to its actual position in the next frame. If, on the other hand, additional information in the form of its coarse velocity is supplied, then the filter will propagate the MB to a location in which the MB is more likely to be detected in the next frame.

We now provide a short description of how OF can be estimated from the low-resolution frames, and how these measurements are incorporated in the Kalman filter framework described in Section II-C to update the state of each localized MB.

The basic assumption in OF estimation is the pixel intensity consistency assumption, or

$$Z_k[x,y] = Z_{k+1}[x + u_{x_k}, y + u_{y_k}], \quad (11)$$

where $[x,y]$ are all pixel coordinates and $[u_{x_k}, u_{y_k}]$ are their corresponding two-dimensional velocities at frame k. Expanding $Z_{k+1}[x+u_{x_k}, y+u_{y_k}]$ to its first order Taylor series around pixel $[x, y]$ yields

$$Z_{k+1}[x+u_{x_k}, y+u_{y_k}] \approx Z_k[x,y] + \mathbf{u}_k^T \boldsymbol{\nabla} Z_k[x,y] + Z_{k_t}[x,y],$$

where $\mathbf{u}_k = [u_{x_k}, u_{y_k}]^T$, $\boldsymbol{\nabla} Z_k[x,y] = [Z_{k_x}[x,y], Z_{k_y}[x,y]]^T$ are the spatial derivatives of $Z_k[x,y]$ and $Z_{k_t}[x,y]$ is the temporal derivative of $Z_k[x,y]$ between frames $k+1$ and k.

Algorithm 2 Triple-SAT

Require: Low-resolution movie $\mathbf{Z}_k, k = 1, \ldots, K$

  while $k \leq K$ **do**

    1: Estimate OF on $\mathbf{z}_k$ using MATLAB's `opticalFlow` command

    2: Perform sparse super-resolution on $\mathbf{z}_k$ using Algorithm 1

    3: reconstruct new measurement vectors for all P localized MBs $\mathbf{y}_k^p, p = 1, \ldots, P$

    4: Associate $\mathbf{y}_k^p$ to existing tracks $\mathbf{T}^p, p = 1, \ldots, P$ / open new tracks / close old tracks using the MHT algorithm

    5: Update last state of existing / new tracks $\mathbf{T}^p$ using (9)

    6: Propagate last state of existing / new tracks $\mathbf{T}^p$ with (7)

    7: Construct weighting matrix $\mathbf{W}_{k+1}$ using (10)

    8: $k \leftarrow k + 1$

  end while

  return Super-resolved frames $\mathbf{i}_k, i = 1, \ldots, K$ and MB tracks

---

Since we perform registration of the images prior to Triple-SAT, the only objects moving in the US sequence are the MBs. Hence, the OF estimation process estimates their velocities on the low resolution grid. By enforcing (11), we obtain the gradient constraint equation

$$\mathbf{u}_k^T \boldsymbol{\nabla} Z_k[x, y] = -Z_{k_t}[x, y]. \tag{12}$$

Denoting

$\mathbf{v}_k = [u_{x_k}[1,1], \ldots, u_{x_k}[M,M], u_{y_k}[1,1], \ldots, u_{y_k}[M,M]]^T$,
$\mathbf{Z}_{k_t} = -[Z_{k_t}[1,1], \ldots, Z_{k_t}[M,M]]^T$,
$\mathbf{A}_{x_k} = \mathrm{diag}[Z_{k_x}[1,1], \ldots, Z_{k_x}[M,M]]$,
$\mathbf{A}_{y_k} = \mathrm{diag}[Z_{k_y}[1,1], \ldots, Z_{k_y}[M,M]]$,

with $\mathbf{D}_k = [\mathbf{A}_{x_k}, \mathbf{A}_{y_k}]$, (12) is written in matrix form as

$$\mathbf{D}_k \mathbf{v}_k = \mathbf{Z}_{k_t}. \tag{13}$$

Equation (13) represents an ill-posed model, since $\mathbf{v}_k$ has $2M^2$ unknowns, but only M^2 measurements are given in $\mathbf{D}_k$.

Many ways to regularize and estimate the OF exist. For example, Horn and Schunk [21] assume global l_2 smoothness of the OF field (12) over all image pixels to estimate the OF and solve

$$\min_{\tilde{\mathbf{u}}_k} ||\mathbf{D}_k \mathbf{v}_k - \mathbf{Z}_{k_t}||_2^2 + \mu ||\boldsymbol{\nabla} \mathbf{u}_k||_2^2,$$

for some regularization parameter $\mu > 0$, where $\boldsymbol{\nabla} \mathbf{u}_k$ are the spatial derivatives of $\mathbf{u}_k$. Lucas and Kanade [23] assume that locally, the OF field is constant and minimize a least-squares criterion,

$$\min_{\mathbf{v}_k} ||\mathbf{M}_k(\tilde{\mathbf{D}}_k \tilde{\mathbf{v}}_k - \mathbf{Z}_{k_t})||_2^2,$$

with $\tilde{\mathbf{v}}_k \in \mathbb{R}^2$ being the lateral and axial velocities,

$$\tilde{\mathbf{D}}_k = \begin{bmatrix} Z_{k_x}[1,1] & Z_{k_y}[1,1] \\ \vdots & \vdots \\ Z_{k_x}[M_\Omega, M_\Omega] & Z_{k_y}[M_\Omega, M_\Omega] \end{bmatrix},$$

$\mathbf{M}_k$ being a window weighting matrix which favors pixels at its center compared with pixels at its periphery and $M_\Omega < M$. The latter minimization problem is solved locally for a small neighborhood on a pixel-wise based manner. OF estimation methods are easily implemented using the `opticalFlow` command in MATLAB. In practice we achieved good performance with the method of Lukas and Kanade with a Gaussian smoothing kernel and a standard deviation of 1.5 pixels.

In practice, each low-resolution frame $\mathbf{Z}_k$ is first interpolated to the size of the $N \times N$ super-resolved images $\mathbf{I}_k$, and OF estimation is performed subsequently. This procedure ensures that each pixel in the super-resolved image is associated with a velocity vector from its corresponding interpolated low-resolution frame. Together, the obtained velocities are considered as measurements for the Kalman filter, along with MBs localizations from the super-resolved frame $\mathbf{I}_k$. Formally, we denote the xy velocity fields obtained by OF estimation over the interpolated low-resolution frame $\mathbf{Z}_k$ as $\mathbf{V}_{x_k}$ and $\mathbf{V}_{y_k}$. That is, both $\mathbf{V}_{x_k}$ and $\mathbf{V}_{y_k}$ are $N \times N$ matrices, and each of their pixels corresponds to the pixel-wise lateral and axial estimated velocities, respectively. Next, for MBs localized in pixels $[i_p, j_p]$ from $\mathbf{I}_k$, we associate the corresponding velocity values from $\mathbf{V}_{x_k}$ and $\mathbf{V}_{y_k}$,

$$\begin{aligned} y_k^p[2] &= V_{x_k}[i_p, j_p], \\ y_k^p[4] &= V_{y_k}[i_p, j_p], \quad p = 1, \ldots, P. \end{aligned} \tag{14}$$

Thus, the first and third entries of the measurement vector $\mathbf{y}_k^p$ in (6) represent the measured position of the pth localized MB in the kth frame, and the second and fourth entries represent its measured velocity.

Note, that we perform OF estimation on the low-resolution movie and not on the super-resolved frames, since we observed in practice that the pixel intensity consistency assumption (11) does not hold on the super-resolved images. This is because a typical super-resolved image looks like the image displayed in Fig. 4, left panel. The enlarged box shows the localizations of four MBs (smoothed only for display purposes). Typically, in the next frame these MBs move considerably, such that (11) does not hold. In contrast, this assumption holds much better on the low-resolution images, due to the spreading of the diffraction-limited echoes from the MBs over several adjacent pixels.

Since the velocities are measured on the low-resolution images, from non-resolved MBs, we cannot expect them to represent the measured velocities of individually resolved MBs. Rather, these velocities represent a coarse, low-resolution estimate of the averaged velocities from the non-resolved MBs. As such, the velocity measurements are weighted with ten-times larger values in the measurement covariance matrix $\mathbf{R}_k^p$ than the position measurements, to represent the inherent inaccuracy of their measurement.

III. MATERIALS AND METHODS

A. Numerical simulations

We simulated a bolus injection into vascular bifurcations over an acquisition period of 105 frames, with frame rate of 10Hz, such that the total acquisition time is 10.5 seconds. Pixel
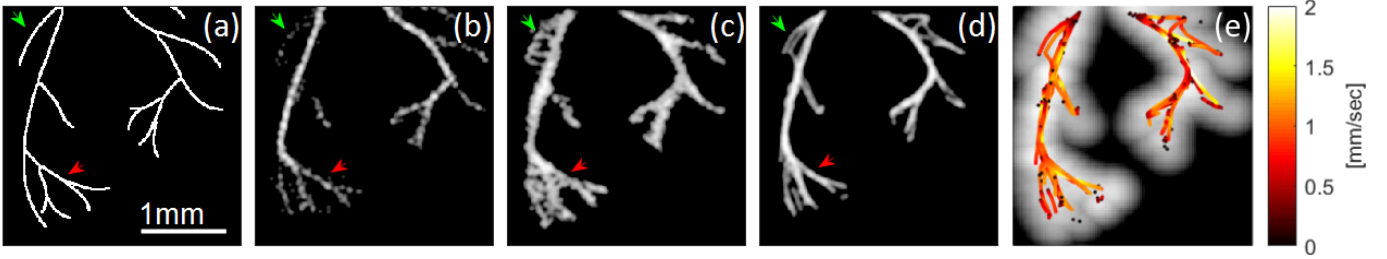
Fig. 6. Simulation results. (a) Ground truth image of bifurcating blood vessels. (b) Super-localization reconstruction. (c) Super-resolution sparse recovery obtained by minimizing (3) via FISTA. (d) Triple-SAT recovery, by accumulating all recovered MB trajectories. (e) Superimposed velocity trajectories over the maximum intensity projection (MIP) image, obtained from the Triple-SAT recovery. All reconstructions are displayed in logarithmic scale, with a dynamic range of 25dB.

size is 0.03×0.03mm^2. The received modulation frequency was 7MHz (second harmonic of 3.5MHz, similar to our *in-vivo* acquisition setup), with a Gaussian PSF having a standard deviation of 0.14mm in the axial direction and 0.16mm in the lateral direction. MB velocities' magnitudes and directions were generated by taking the maximum between values drawn from a normal distribution with a mean of 1mm/sec and standard deviation of 1mm/sec, and zero.

We set $P = 3$ and recover the super-resolved images on a three times denser grid than the low-resolution grid with $\lambda = 0.035$. We iterate over 4000 iterations per frame and set $\epsilon = 0.01$, $\rho = 300$, and $\mathbf{R}_k = \text{diag}\{0.1, 1, 0.1, 1\}$. In the MHT algorithm, we set the probability for not detecting an existing target as 0.1, the probability for a new target to appear as 0.1, the probability for false-alarm as 0.01 and a maximum number of leaves of 6.

B. In-vivo experiments

CEUS data of human prostates was acquired at the AMC university hospital (Amsterdam, the Netherlands), using a transrectal ultrasound probe and an iU22 scanner (Philips Healthcare, Bothell, WA) operating in a contrast-specific mode at a frame rate of $1/\Delta T = 10$Hz. A 2.4-mL MB bolus of Sonovue$^{\text{TM}}$ (Bracco, Milan, Italy) was administered intravenously, and 100-150 frames (10-15 seconds) were collected for further analysis. Pixel size is 0.146×0.146mm^2.

We consider two examples taken from the *in-vivo* scan. For all experiments, we set $P = 4$ and recover the super-resolved images on a four times denser grid than the low-resolution grid. In both cases we use 2000 iterations per frame, $\epsilon = 0.5$, $\rho = 500$, and $\mathbf{R}_k = \text{diag}\{0.1, 1, 0.1, 1\}$. For both examples, in the MHT algorithm, we set the probability for not detecting an existing target as 0.1, the probability for a new target to appear as 0.5, the probability for false-alarm as 0.01 and a maximum number of leaves of 6.

IV. RESULTS

A. Numerical simulations

Figure 6 shows reconstruction results of the simulated dataset of flowing MBs within a simulated vascular network. Panel (a) shows the ground truth architecture, while panels (a) (b) and (c) show the reconstruction results of standard super-localization [41], sparsity-driven super-resolution [16]

(minimizing (3)) and Triple-SAT, respectively. Panel (e) shows an overlay of MB sub-diffraction trajectories, colored by their estimated velocities over the maximum intensity projection image (MIP) image.

We observe that the Triple-SAT recovery (panel (d)) seems the smoothest and most continuous, compared with the super-localization and sparsity-based reconstructions, depicted in panels (b) and (c), respectively. The green arrows in the panels show a protruding blood vessel, which is almost absent in panel (b), and in panel (c) it seems connected to an additional vessel. In contrast, in panel (d), this blood vessel is more clearly depicted. The red arrows indicate clear bifurcations in the Triple-SAT reconstruction (panel d), which are absent in the sparsity-based recovery (panel c). This area is more clearly reconstructed in the super-localization image (panel d) than in the sparsity-driven recovery, but seems less smooth and clear compared with the Triple-SAT image. Though Triple-SAT shows some reconstruction artifacts, it depicts the clearest and smoothest image compared to other methods.

Additionally, panel (e) reveals that the estimated velocities are in the range $0 - 2$mm/sec, as expected from the simulation. A histogram of the measured velocities is given in panel (c) of Fig. 9, where the velocities distribution is indeed between $0 - 2$mm/sec.

B. In-vivo experiments

In this section, we present *in-vivo* reconstruction results of Triple-SAT. Ultrasound acquisition parameters and reconstruction parameters are given in Section III. Figures 7 and 8 compare between different reconstructions in two areas of a prostate CEUS scan. In both figures, panel (a) shows the MIP image. This image is diffraction limited and was generated as reference for standard non-super-resolution image processing by taking the pixel-wise maximum value over the entire movie. Panel (b) shows sparsity-based super-resolution as obtained by minimizing (3) via FISTA, while panel (c) shows the Triple-SAT image. Lastly, panel (d) displays an overlay of the estimated velocities' trajectories on the MIP image.

Comparing panels (a) to (b) and (c) of both figures, we notice how sparse recovery (panel (b)) achieves a super-resolved depiction of the fine vasculature. However, the Triple-SAT image (panel (c)) qualitatively appears more smooth and continuous, showing distinct trajectories which are absent in
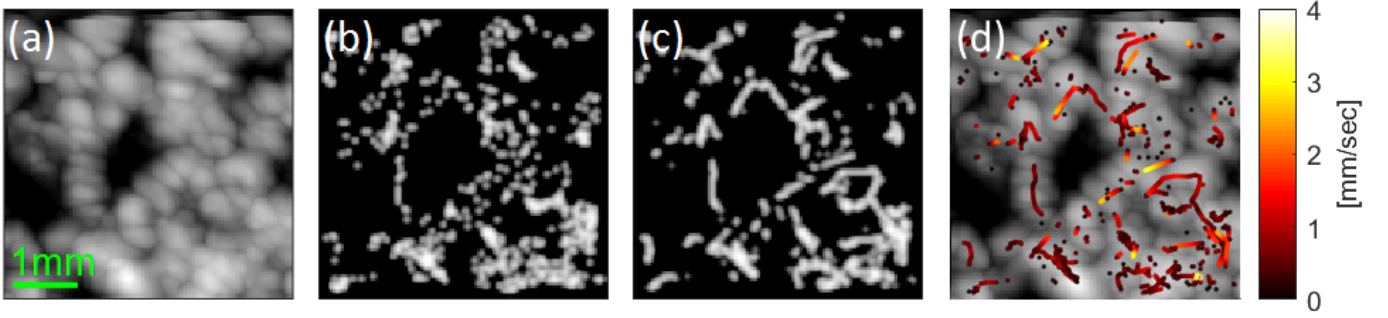
Fig. 7. Triple-SAT applied to an *in-vivo* scan from a human-prostate. (a) MIP image from 100 frames. (b) Super-resolution sparse recovery obtained by minimizing (3) via FISTA. (c) Triple-SAT recovery, by accumulating all recovered MB trajectories. (d) Superimposed velocity trajectories over the MIP image, obtained from the Triple-SAT recovery. All images are displayed in logarithmic scale, with a dynamic range of 30dB.
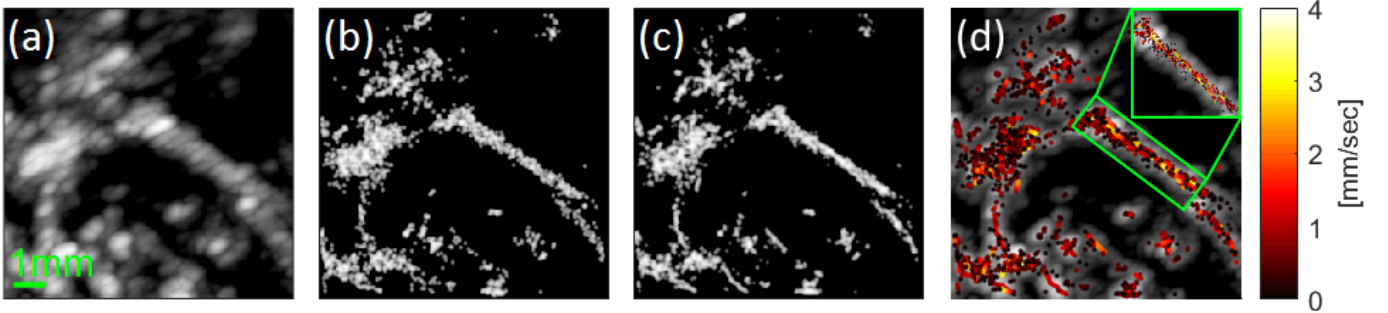

Fig. 8. Additional example of Triple-SAT recovery of an *in-vivo*, human-prostate scan. (a) MIP image from 150 frames. (b) Super-resolution sparse recovery obtained by minimizing (3) via FISTA. (c) Triple-SAT recovery by accumulating all recovered MB trajectories. (d) Superimposed velocity trajectories over the MIP image, obtained from the Triple-SAT recovery. All images are displayed in logarithmic scale, with a dynamic range of 30dB.
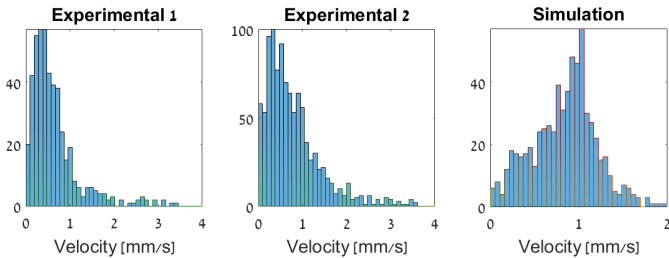

Fig. 9. Estimated velocity histograms. (a) Experimental dataset in Fig. 7. (b) Experimental dataset in Fig. 8. (c) Simulation dataset in Fig. 6.

the sparse recovery images. Finally, panels (d) in both figures present velocity magnitude estimations from the tracked MBs. The plotted trajectories are sub-diffraction in size, which is clearly noticeable for example in panel (d) of Fig. 7. The green box in the upper-right corner depicts an enlarged region, were clear MB trajectories with velocities of $\sim 2 - 3$mm/sec are observed. The vast majority of obtained MB flow velocities are on the order of up to $1 - 2$mm/sec, in line with previous observations on blood flow in micro-vessels [13]. This is also confirmed by the velocity magnitude histograms, displayed in Fig. 9.

Panels (d) in both figures depict many point-like and short trajectories, with low velocities, alongside longer and smoother trajectories. We attribute these point-like trajectories to the fact that the vascular bed of the prostate is inherently 3D, with many blood vessels crossing the imaging plane of the probe.

Thus, lateral and axial velocities (with respect to the transducer position) of MBs flowing within these blood vessels can be small. In contrast, the simulated MBs in Fig. 6 are simulated in a plane, and clearly show long and smooth trajectories.

V. DISCUSSION AND CONCLUSIONS

This is the first work to exploit the inherent motion kinematics of individual MBs as a structural prior for super-resolution. Since individual MBs flow within blood vessels, their positions can be predicted from one frame to the next. In Triple-SAT we exploit this additional information to improve sparse-recovery, by solving a support aware minimization problem, as formulated in (4). Using Kalman filtering, we track and propagate the trajectories of individual MBs from one frame to the next. Moreover, we introduce velocity measurements via optical flow estimation to improve the tracking process for super-resolution imaging. Figures 7 and 8 show the power and potential of Triple-SAT on *in-vivo* data. Both clear and smooth super-resolution imaging are achieved, as well as a quantitative measurement of the flow velocities of individual MBs.

As with other sparsity-based super-resolution methods, Triple-SAT operates well with high MB concentrations, for which significant MB overlap is present. This is a key observation in bridging the gap between low-rate clinical scanners with a fixed frame-rate, and clinical MBs doses. We note that Triple-SAT can also work with even higher frame-rates, such as fast plane-wave scans. This, for example, would allow

to produce super-resolved images in approximately $\leq$ 10ms using $100 - 150$ frames acquired with a frame-rate of 10KHz. By exploiting the sparse nature of the individual MB echoes, Triple-SAT is able to depict the vasculature with a relatively low number of frames. In our *in-vivo* experiments we used two datasets of 100 and 150 frames. However, as MB density increases even further, several mechanisms of Triple-SAT might fail. First, the sparse-recovery algorithm may not be able to accurately detect and localize all of the MBs in each frame. Second, MHT data-to-track association may also fail to properly associate new localizations to existing tracks, as resolved MBs become extremely close to one another. Lastly, OF estimation will fail to produce reliable results in areas of many overlapping MBs which move in different directions, resulting in an almost-zero averaged velocity estimate on the low-resolution grid. Reducing the frame-rate even more will also cause OF estimation and MHT data-to-track association to fail. The former, since the basic assumption of pixel intensity consistency breaks down as the frame-rate decreases. The latter, since the association of new measurements to existing tracks becomes less likely than the opening of new tracks, even if they belong to previous tracks. Yet, as we report in this work, for clinical bolus doses and 10Hz scanners, Triple-SAT reliably recovers the vascular bed.

In [15], a similar model to (1) was introduced over the beam-formed complex IQ signal, while here we assume such a model over the real-valued intensity images. Nonetheless, in practice, both in simulations and *in-vivo* experiments, we observed good reconstruction performance using model (1), as we presented in Section IV.

The use of a Kalman filter for MB tracking has two main motivations. First, the Kalman filter is an online estimator, which is suitable for real-time applications. Using this filter for online tracking of MBs can lead to a real-time clinical application of Triple-SAT. Second, it is known that capillary flow is non-turbulent [8, 19]. Thus, a simple linear propagation model can be used for the tracking procedure.

There are two main limitations to Triple-SAT. The first is inherent to all ultrasound super-resolution techniques. MBs must flow through the vasculature in order to detect it, thus setting a minimal acquisition time for any super-resolution imaging technique. In Triple-SAT, by using high-concentration bolus doses and sparse-recovery, we reduce the acquisition time, but only to the degree that MBs flow within the finest blood vessels during that period. Second, Triple-SAT includes several parameters which should be selected properly, among which are the sparsity regularization parameter λ, ϵ, ρ and the probabilities for the MHT algorithm. In this work, these parameters were chosen manually, and calibrated according to the simulation, but additional validation and testing of their selection is required.

Before concluding, we would like to discuss some computational aspects of online sparse tracking, as the number of detected MBs grows. The authors of [29] suggested an l_1 relaxed adaptation of the Kalman filter to account for the possible exponential growth in computational complexity with the problem dimensions. In practice, although in Triple-SAT we apply a Kalman filter to each localized MB, this computational growth was not observed to be dramatic, even when tens of MBs were tracked simultaneously. We ascribe this to the fact that the state of each MB is relatively low dimensional (four entries of positions, and velocities). As such, matrix inversions are relatively inexpensive.

Another possible computational burden stems from the MHT algorithm, which is known to grow exponentially in complexity as the number of tracks increases. In [8], the authors considered a modified version of Markov Chain Monte Carlo (MCMC) data association [42] to account for this growth. The computational complexity of MHT can also be controlled by limiting the pruning depth, achieving a trade-off between accurate data association and computational complexity. In general, any automatic association algorithm may be used in the uppermost block of Fig. 3 instead of MHT, such as the joint probabilistic data association (JPDA) [43] or the MCMC algorithm of [8].

To conclude, in this work, we presented a new algorithm to improve sparsity-based super-resolution CEUS imaging. By formulating a weighted sparse recovery minimization problem, combined with on-line tracking of individual MBs, we are able to improve the sparse recovery process. With Triple-SAT we achieve a smoother depiction of the vasculature as well as provide quantitative information regarding MB kinematics. We applied our algorithm to both simulations and *in-vivo* human prostate scans, obtained from low frame-rate, clinically-approved US machines, demonstrating super-resolution recovery of the vascular bed with 100-150, high-MB-density frames. Since Triple-SAT employs an on-line estimation process, it may be suitable for real-time applications within commercially available US machines.

REFERENCES

[1] R. Schlief, "Ultrasound contrast agents." *Current opinion in radiology*, vol. 3, no. 2, pp. 198–207, 1991.

[2] N. De Jong, F. Ten Cate, C. Lancee, J. Roelandt, and N. Bom, "Principles and recent developments in ultrasound contrast agents," *Ultrasonics*, vol. 29, no. 4, pp. 324–330, 1991.

[3] D. Cosgrove, "Ultrasound contrast agents: an overview," *European journal of radiology*, vol. 60, no. 3, pp. 324–330, 2006.

[4] J. M. Hudson, R. Williams, C. Tremblay-Darveau, P. S. Sheeran, L. Milot, G. A. Bjarnason, and P. N. Burns, "Dynamic contrast enhanced ultrasound for therapy monitoring," *European journal of radiology*, vol. 84, no. 9, pp. 1650–1657, 2015.

[5] C. Errico, J. Pierre, S. Pezet, Y. Desailly, Z. Lenkei, O. Couture, and M. Tanter, "Ultrafast ultrasound localization microscopy for deep super-resolution vascular imaging," *Nature*, vol. 527, no. 7579, pp. 499–502, 2015.

[6] M. A. OReilly and K. Hynynen, "A super-resolution ultrasound method for brain vascular mapping," *Medical physics*, vol. 40, no. 11, 2013.

[7] K. Christensen-Jeffries, R. J. Browning, M.-X. Tang, C. Dunsby, and R. J. Eckersley, "In vivo acoustic super-resolution and super-resolved velocity mapping using microbubbles," *IEEE transactions on medical imaging*, vol. 34, no. 2, pp. 433–440, 2015.

[8] D. Ackermann and G. Schmitz, "Detection and tracking of multiple microbubbles in ultrasound b-mode images," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 63, no. 1, pp. 72–82, 2016.

[9] J. Foiret, H. Zhang, T. Ilovitsh, L. Mahakian, S. Tam, and K. W. Ferrara, "Ultrasound localization microscopy to image

and assess microvasculature in a rat kidney," *Scientific reports*, vol. 7, no. 1, p. 13662, 2017.

[10] A. Bar-Zion, C. Tremblay-Darveau, O. Solomon, D. Adam, and Y. C. Eldar, "Fast vascular ultrasound imaging with enhanced spatial resolution and background rejection," *IEEE transactions on medical imaging*, vol. 36, no. 1, pp. 169–180, 2017.

[11] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, "Imaging intracellular fluorescent proteins at nanometer resolution." *Science (New York, N.Y.)*, vol. 313, no. 5793, pp. 1642–1645, 2006.

[12] M. J. Rust, M. Bates, and X. Zhuang, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)." *Nature methods*, vol. 3, no. 10, pp. 793–795, 2006.

[13] R. J. van Sloun, L. Demi, A. W. Postema, J. J. de la Rosette, H. Wijkstra, and M. Mischi, "Ultrasound-contrast-agent dispersion and velocity imaging for prostate cancer localization," *Medical image analysis*, vol. 35, pp. 610–619, 2017.

[14] Y. C. Eldar, *Sampling Theory: Beyond Bandlimited Systems*. Cambridge University Press, 2015.

[15] A. Bar-Zion, O. Solomon, C. Tremblay-Darveau, D. Adam, and Y. C. Eldar, "Sparsity-based ultrasound super-resolution hemodynamic imaging," *arXiv preprint arXiv:1712.00648*, 2017.

[16] R. J. van Sloun, O. Solomon, Y. C. Eldar, H. Wijkstra, and M. Mischi, "Sparsity-driven super-resolution in clinical contrast-enhanced ultrasound," in *Ultrasonics Symposium (IUS), 2017 IEEE International*. IEEE, 2017, pp. 1–4.

[17] A. Bar-Zion, O. Solomon, C. Tremblay-Darveau, D. Adam, and Y. C. Eldar, "Sparsity-based ultrasound super-resolution imaging," in *Proceedings of the 23rd European symposium on Ultrasound Contrast Imaging*, 2017, pp. 156–157.

[18] R. J. van Sloun, L. Demi, A. W. Postema, J. J. De La Rosette, H. Wijkstra, and M. Mischi, "Entropy of ultrasound-contrast-agent velocity fields for angiogenesis imaging in prostate cancer," *IEEE transactions on medical imaging*, vol. 36, no. 3, pp. 826–837, 2017.

[19] J. A. Jensen, *Estimation of blood velocities using ultrasound: a signal processing approach*. Cambridge University Press, 1996.

[20] J. L. Barron, D. J. Fleet, and S. S. Beauchemin, "Performance of optical flow techniques," *International journal of computer vision*, vol. 12, no. 1, pp. 43–77, 1994.

[21] B. K. Horn and B. G. Schunck, "Determining optical flow," *Artificial intelligence*, vol. 17, no. 1-3, pp. 185–203, 1981.

[22] S. Baker and I. Matthews, "Lucas-kanade 20 years on: A unifying framework," *International journal of computer vision*, vol. 56, no. 3, pp. 221–255, 2004.

[23] B. D. Lucas, T. Kanade *et al.*, "An iterative image registration technique with an application to stereo vision," 1981.

[24] R. E. Kalman *et al.*, "A new approach to linear filtering and prediction problems," *Journal of basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[25] Y. Bar-Shalom and X.-R. Li, *Multitarget-multisensor tracking: principles and techniques*. YBs London, UK:, 1995, vol. 19.

[26] G. Ascenti, M. Gaeta, C. Magno, S. Mazziotti, A. Blandino, D. Melloni, and G. Zimbaro, "Contrast-enhanced second-harmonic sonography in the detection of pseudocapsule in renal cell carcinoma," *American Journal of Roentgenology*, vol. 182, no. 6, pp. 1525–1530, 2004.

[27] N. Vaswani and W. Lu, "Modified-cs: Modifying compressive sensing for problems with partially known support," *IEEE Transactions on Signal Processing*, vol. 58, no. 9, pp. 4595–4607, 2010.

[28] N. Vaswani, "Kalman filtered compressed sensing," in *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*. IEEE, 2008, pp. 893–896.

[29] D. Angelosante, S. I. Roumeliotis, and G. B. Giannakis, "Lasso-kalman smoother for tracking sparse signals," in *Signals, Systems and Computers, 2009 Conference Record of the Forty-Third Asilomar Conference on*. IEEE, 2009, pp. 181–185.

[30] D. Angelosante, G. B. Giannakis, and E. Grossi, "Compressed sensing of time-varying signals," in *Digital Signal Processing, 2009 16th International Conference on*. IEEE, 2009, pp. 1–8.

[31] O. Solomon, Y. C. Eldar, M. Mutzafi, and M. Segev, "Sparcom: Sparsity based super-resolution correlation microscopy," *arXiv preprint arXiv:1707.09255*, 2017.

[32] A. Beck and M. Teboulle, "A Fast Iterative Shrinkage-Thresholding Algorithm," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[33] Y. C. Eldar and G. Kutyniok, *Compressed sensing: theory and applications*. Cambridge University Press, 2012.

[34] E. J. Candes, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted 1 minimization," *Journal of Fourier analysis and applications*, vol. 14, no. 5-6, pp. 877–905, 2008.

[35] D. Reid, "An algorithm for tracking multiple targets," *IEEE transactions on Automatic Control*, vol. 24, no. 6, pp. 843–854, 1979.

[36] C. Kim, F. Li, A. Ciptadi, and J. M. Rehg, "Multiple hypothesis tracking revisited," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 4696–4704.

[37] P. M. Pardalos and N. Desai, "An algorithm for finding a maximum weighted independent set in an arbitrary graph," *International Journal of Computer Mathematics*, vol. 38, no. 3-4, pp. 163–175, 1991.

[38] D. M. Antunes, D. M. de Matos, and J. Gaspar, "A library for implementing the multiple hypothesis tracking algorithm," *arXiv preprint arXiv:1106.2263*, 2011.

[39] D. M. Antunes, D. Figueira, D. M. Matos, A. Bernardino, and J. Gaspar, "Multiple hypothesis tracking in camera networks," in *Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on*. IEEE, 2011, pp. 367–374.

[40] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with applications to tracking and navigation: theory algorithms and software*. John Wiley & Sons, 2004.

[41] M. Ovesnỳ, P. Křížek, J. Borkovec, Z. Švindrych, and G. M. Hagen, "Thunderstorm: a comprehensive imagej plug-in for palm and storm data analysis and super-resolution imaging," *Bioinformatics*, vol. 30, no. 16, pp. 2389–2390, 2014.

[42] S. Oh, S. Russell, and S. Sastry, "Markov chain monte carlo data association for multi-target tracking," *IEEE Transactions on Automatic Control*, vol. 54, no. 3, pp. 481–497, 2009.

[43] T. E. Fortmann, Y. Bar-Shalom, and M. Scheffe, "Multi-target tracking using joint probabilistic data association," in *Decision and Control including the Symposium on Adaptive Processes, 1980 19th IEEE Conference on*. IEEE, 1980, pp. 807–812.