Exploiting flow dynamics for superresolution in contrast-enhanced ultrasound

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Exploiting Flow Dynamics for Superresolution 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 (US) localization microscopy offers new radiation-free diagnostic tools for vascular imaging deep within the tissue. Sequential localization of echoes returned from inert microbubbles (MBs) with low concentration within the blood-stream reveals the vasculature with capillary resolution. Despite its high spatial resolution, low MB concentrations dictate the acquisition of tens of thousands of images, over the course of several seconds to tens of seconds, to produce a single superresolved image. Such long acquisition times and stringent constraints on MB 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 MB overlap. Here, we further improve the spatial resolution and visual vascular reconstruction quality of sparsity-based superresolution US imaging from low-frame rate acquisitions, by exploiting the inherent flow of MBs and utilize their motion kinematics. We also provide quantitative measurements of MB velocities and show that our approach achieves higher MB recall rate than the state-of-the-art techniques, while increasing contrast agents concentration. Our method relies on simultaneous tracking and sparsity-based detection of individual MBs in a frame-by-frame manner, and as such, may be suitable for real-time implementation. The effectiveness of the proposed approach is demonstrated on both simulations and an *in vivo* contrast-enhanced human prostate scan, acquired with a clinically approved scanner operating at a 10-Hz frame rate.

Index Terms—Compressed sensing, contrast agents, Kalman filter, superresolution, ultrasound (US).

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)

O. Solomon is with the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa 3200003, Israel (e-mail: orensol@campus.technion.ac.il).
R. J. G. van Sloun and M. Mischi are with the Department of Electrical Engineering, Eindhoven University of Technology, 5612 AZ Eindhoven, The Netherlands (e-mail: r.j.g.v.sloun@tue.nl; m.mischi@tue.nl).
H. Wijkstra is with the Department of Electrical Engineering, Eindhoven University of Technology, 5612 AZ Eindhoven, The Netherlands, and also with the Department of Urology, Academic Medical Center, University of Amsterdam, 1012 WX Amsterdam, The Netherlands (e-mail: h.wijkstra@tue.nl).
Y. C. Eldar is with the Faculty of Math and CS, Weizmann Institute of Science, Rehovot 7610001, Israel (e-mail: yonina@weizmann.ac.il).

imaging takes the advantage of inert MBs that are injected into the bloodstream, as means to image blood vessels with the improved contrast, compared with the standard B-mode US imaging [4]. In recent years, superresolution US imaging emerged and enabled the fine visualization and detailed assessment of capillary blood vessels *in vivo* [5]–[11]. This approach was also extended to 3-D imaging for the identification of microvascular morphology features of tumor angiogenesis [12]. Superresolution US relies on concepts borrowed from superresolution fluorescence microscopy techniques such as photo-activated localization microscopy (PALM) and stochastic optical reconstruction microscopy (STORM) [13], [14], which localize individual fluorescing molecules with subpixel precision over many frames and sum all localizations to produce a superresolved 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 superresolved image of the vascular bed with a tenfold 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 pinpointing the centroid of a single Gaussian for each detected MB in the captured movie.

Despite yielding a substantial improvement in spatial resolution, superresolution US imaging typically requires tens of thousands of images to produce a single superresolved 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 [15]. These limitations hamper the clinical applicability of localization-based superresolution techniques.

To overcome the temporal limitation of localization-based superresolution without compromising the spatial resolution of the reconstructed image, sparsity-based [16] approaches were recently proposed [17]–[20]. These techniques favor overlapping MBs to reduce the total acquisition time. Thus, sparsity-based methods achieve faster temporal resolution using standard clinical concentrations of MBs. In [17]–[19], sparsity-based superresolution US hemodynamic imaging (SUSHI), using ultrafast plane-wave acquisition, demonstrated a superresolved time-lapse movie of 25 Hz, showing superresolved hemodynamic changes in blood flow within a rabbit's kidney. In [20], using a clinically approved scanner with an acquisition rate of 10 Hz, a superresolved image of a human

prostate vasculature was shown. In particular, 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 superresolved image.

One major difference between superresolution in US and in microscopy is that the point emitters in US are flowing inside the blood vessels, whereas in microscopy superresolution imaging, fluorescent molecules are fixed to the subcellular organelles. Since the motion of individual MBs is not random but rather within blood vessels, this can be exploited to improve the recovery process. This paper builds on our previous results on superresolution US imaging [19], [20] by exploiting the flow kinematics of individual MBs as an additional prior in the sparse recovery process.

While previous superresolution works focused on ultrafast plane-wave image acquisition, see [5], [17], [19], most clinically used scanners are low-rate scanners (10–25 Hz). When using high frame-rate scanners, e.g., ultrafast plane-wave imaging, fast superresolution imaging can be achieved via SUSHI [17]–[19], exhibiting a smooth depiction of the superresolved vessels, with a relatively low-complexity algorithm. However, as the frame-rate decreases, MB detections become more sporadic, even when using SUSHI, resulting in an inconsistent depiction of the vessels. Thus, the so-called missing information needs to be filled-in by other means, albeit with higher computational cost.

Here, our aim is to bridge the gap between superresolution 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 demonstrated 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 [21], 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. Each MB track is used to estimate the position of the MBs and fill-in for the missing spatial information due to low-rate scanning, thus providing a smooth depiction of the superresolved vessels. Since the capillary flow is nonturbulent (peak Reynolds number of 0.001) [8], [22], 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. Our method is titled simultaneous sparsity-based superresolution and tracking, or 3SAT[1] (pronounced triple-SAT). Since our approach tracks individual MBs, it is possible to also estimate
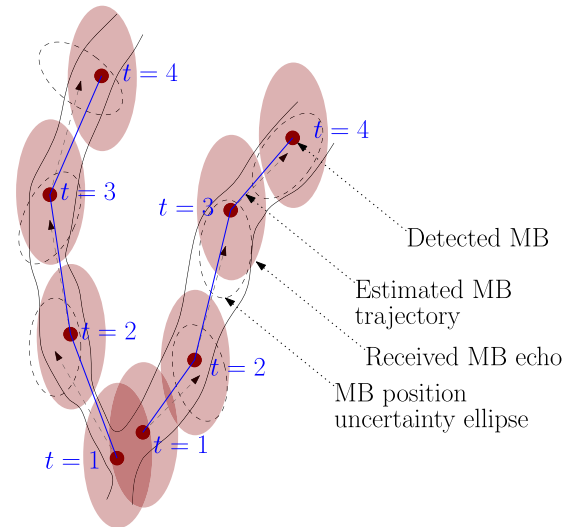
Fig. 1. Proposed concept of 3SAT. Individual MBs flow within a blood vessel depicted here as a bifurcation by black solid contour lines. Large, transparent, and red ellipses represent the echoes measured from individual MBs. In frame $t = 1$, MBs are detected using sparse recovery (small red dots). Applying a Kalman filter, their positions are propagated to the next frame ($t = 2$) as marked by the black dashed arrows. Using the error covariance matrix of the filter, the ellipses of most likely positions for the MBs are generated as illustrated by the black dashed ellipses. These ellipses are then used as weights in the sparse recovery process in the next frame ($t = 2$), and so on. Blue lines: estimated trajectories of the MBs.

their velocities. We provide *in vivo* superresolution CEUS imaging of a human prostate and show that its velocity estimation agrees with the previously published results [15]. An illustration of the proposed concept is shown 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, 3SAT 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. Instead, 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, 3SAT exploits detections from the previous frames to improve the detections in the next frame, using sparse recovery to overcome MBs overlap, resulting from clinical bolus doses. Second, the coarse measurements of MB movements based on optical flow (OF) estimation [23]–[26] over the captured low-resolution sequence are exploited to improve the tracking performance. Thus, 3SAT incorporates not only position measurements but also velocity measurements in the adopted Kalman filtering framework [27], [28]. 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 the Kalman filtering, as a mean to include additional information from the previous frames and improve the overall estimation accuracy.

The rest of this paper is organized as follows. Section II describes 3SAT and each of its building blocks. Sections III
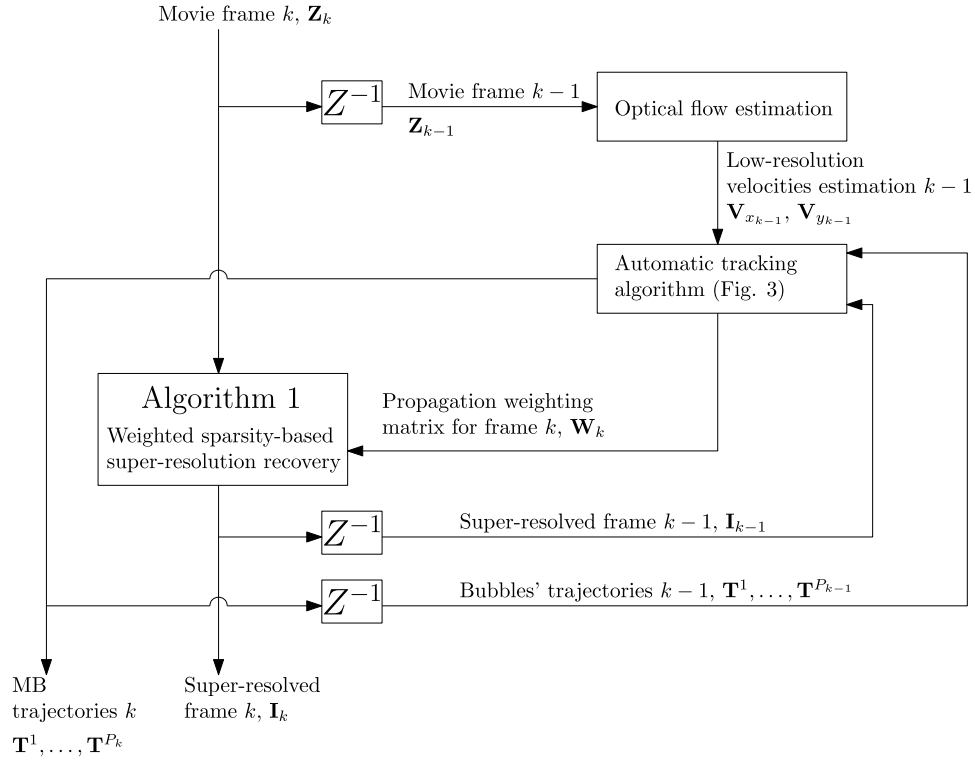
Fig. 2. Main building blocks of 3SAT. First, MBs velocities are estimated from frame k using OF estimation. The detected MBs from the previous frame $k-1$ are then propagated to frame k, assuming a constant velocity model using the Kalman filter and the measured velocities obtained from the OF estimation. This procedure yields a new estimate for the true MB positions and velocities, and also forms the weighting matrix for frame k. This weighting matrix is then used in the sparse recovery process to yield the superresolved frame k. This process repeats itself for each new frame in the movie. Here, Z^{-1} represents a delay of one time unit. The entire 3SAT process is summarized in Algorithm 2.

and IV present *in silico* as well as *in vivo* results. A discussion and conclusions are provided in Section V.

Throughout this paper, x represents a scalar, $\mathbf{x}$ represents a vector, $\mathbf{X}$ represents a matrix, and $\mathbf{I}_{N \times N}$ is the $N \times N$ identity matrix. The notation $\mathbf{X}^T$ represents the transpose of $\mathbf{X}$ and $\mathbf{X}^H$ its conjugate transpose. We denote by $||\cdot||_p$ the standard p-norm and by $||\cdot||_F$ the Frobenius norm. Subscript x_l denotes the lth element of $\mathbf{x}$ and $\mathbf{x}_l$ is the lth column of $\mathbf{X}$, while superscript $(\cdot)^p$ refers to the pth MB. The estimated vector in frame k, given the estimate in the $(k-1)$th frame, is indicated by $\mathbf{s}_{k|k-1}$. Similarly, $\mathbf{P}_{k|k-1}$ indicates its estimated covariance matrix k, given the $k-1$ estimate. The ijth element of a matrix $\mathbf{A}$ is denoted $A[i, j]$. The notation $\tilde{\mathbf{x}}$ explicitly indicates that x is the Fourier domain representation of $\mathbf{x}$.

II. SIMULTANEOUS SPARSITY-BASED SUPERRESOLUTION AND TRACKING

A. Principle

In this work, our primary aim is to improve sparsity-based superresolution US from movies that are acquired from low-rate clinical scanners, where we have access only to the final intensity images displayed on the screen. We start from a CEUS 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 the signal from MBs [15], [21], [29], such that only MBs are visualized. Since our primary focus is on sequences acquired from low-rate scanners during the

acquisition period movement of the probe and scanned organ are inevitable. This movement introduces inaccuracies in the estimation process of MB positions and velocities [10]. Thus, prior to 3SAT processing, all frames are registered, as described in [20], to compensate for this inaccuracy. After tissue and MBs separation, the spatially correlated tissue sequence is used for registration. For each image in the sequence, an affine transformation is determined which maps it to the first image in the sequence. This transformation is then applied to each corresponding contrast image so that at the end, all contrast images are spatially aligned. The registration process is performed in MATLAB (The MathWorks, Inc.) using the imregtform function. After registration, the input data for 3SAT consist of K registered low-resolution frames.

Fig. 2 shows the main flow and building blocks of 3SAT. Given the weighting matrix, based on trajectories estimated from the $(k-1)$th frame, 3SAT performs weighted sparse recovery to estimate the positions of the MBs on a high-resolution grid in the kth frame. Next, we estimate 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, both positions and corresponding velocity measurements are obtained, which are used in the automatic tracking algorithm to update the positions and velocities of the individually detected MBs via 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 3SAT 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 [30]–[33]. 3SAT differs from these methods in the following manner. First, previous works assume a first-order recursion for the propagation model of the nonzero 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 the 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, an extended model is considered, which includes the position estimation of the MBs together with their velocities, as presented in Section II-C. Second, as MBs flow over time, new MBs emerge and some MBs vanish from the imaging plane, due to the 3-D 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 3SAT by the use of an automatic association algorithm (as described in Section II-C), combined with Kalman filtering.

We next detail the main building blocks of 3SAT.

B. Weighted Sparse Recovery

We start with a description of our sparse recovery algorithm, which is performed on each frame separately. Similar to [19], a frame is modeled as a summation of L_k individual MB echoes

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

where $u(\cdot, \cdot)$ is the (intensity-based) point spread function (PSF) of the transducer and σ_i is the magnitude of the returned echo from the ith MB located at position (x_i, y_i). The PSF of the transducer is assumed to be known. In practice, the PSF can be measured from the acquired images, as described in [19] and [20].

Following similar derivations to [19], 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 $P \geq 1$ and denote the (vectorized) superresolved 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 superresolved 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 [19], [34] and consider recovering $\mathbf{i}_k$ in the discrete Fourier domain. In this domain, $\tilde{\mathbf{H}}$ has the following structure:

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

where $\otimes$ symbolizes the Kronecker product of two matrices. Here, $\tilde{\mathbf{U}}$ is an $M^2 \times M^2$ diagonal matrix, whose diagonal contains the vectorized 2-D discrete Fourier transform (DFT)

Algorithm 1 FISTA for Minimizing (4)

Require: $\tilde{\mathbf{z}}_k, \tilde{\mathbf{H}}, \mathbf{W}_k, \lambda > 0$, maximum iterations $D_{\max}$
   Initialize $\mathbf{y}_1 = \mathbf{x}_0 = \mathbf{0}$, $\mathbf{w} = \text{diag}\{\mathbf{W}_k\}$, $t_1 = 1$ and $d = 1$
   Calculate $L_f = ||\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}||_2$
   while $d \leq D_{\max}$ or stopping criteria not fulfilled **do**
      $\mathbf{g}_d = \tilde{\mathbf{H}}^H\tilde{\mathbf{H}}\mathbf{y}_d - \tilde{\mathbf{H}}^H\tilde{\mathbf{z}}_k$
      $\mathbf{x}_d = \mathcal{T}_{\frac{\lambda}{L_f}\mathbf{w}}\left(\mathbf{y}_d - \frac{1}{L_f}\mathbf{g}_d\right)$
      Project to the non-negative orthant $\mathbf{x}_d(\mathbf{x}_d < \mathbf{0}) = \mathbf{0}$ and to the real numbers $\mathbf{x}_d = \text{real}(\mathbf{x}_d)$
      $t_{d+1} = \frac{1}{2}\left(1 + \sqrt{1 + 4t_d^2}\right)$
      $\mathbf{y}_{d+1} = \mathbf{x}_d + \frac{t_d - 1}{t_{d+1}}(\mathbf{x}_d - \mathbf{x}_{d-1})$
      $d \leftarrow d + 1$
   end while
   return $\mathbf{i}_k = \mathbf{x}_{D_{\max}}$

---

of the PSF, and $\mathbf{F}_M$ is a partial $M \times N$ 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 [34]. The estimation of $\mathbf{i}_k$ is achieved by solving the following convex optimization problem:

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

where $\lambda \geq 0$ is a regularization parameter and $\tilde{\mathbf{z}}_k$ is the Fourier transform of $\mathbf{z}_k$. Note the $\mathbf{i}_k$ is a real quantity.

In [20], the superresolved image is 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:

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

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, (4) is minimized using the fast iterative/shrinkage thresholding algorithm (FISTA) [35], [36], described in Algorithm 1, or by using the reweighted iterative l_1 method [37] (we project to the real numbers to avoid residual imaginary errors in the estimation of $\mathbf{x}_d$, as $\mathbf{i}_k$ is a real quantity). Algorithm 1 is applied 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 $\mathbf{H}^H\mathbf{H}$, 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 the vectors of the same length.

We next describe how to construct $\mathbf{W}_k$ per frame using MBs trajectories, Kalman filtering, and OF.

C. Microbubble Tracking

The (diagonal of the) weighting matrix $\mathbf{W}_k$ is inversely proportional to the accumulated probability of detected 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 shown in Section II-D. We now turn to explain this process. First, the state of the pth MB in frame k is defined as $\mathbf{s}_k^p \in \mathbb{R}^4$ with

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

Here, x_k^p and y_k^p are the Cartesian coordinates that 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, consider an arbitrary frame, k. At this stage, we posses all the states of P_{k-1} previously tracked MBs, $\mathbf{s}_{k-1}^1, \ldots, \mathbf{s}_{k-1}^{P_{k-1}}$. Given the next low-resolution frame $\mathbf{z}_k$, our main two goals are the following.

1) Recover the locations of the L_k MBs which are embodied in frame $\mathbf{z}_k$. The number L_k of MBs in frame k is generally different than the number of MBs in the previous frame L_{k-1}. This possible discrepancy occurs since blood vessels have a 3-D topology, and consequently MBs may shift in and out of the imaging plane.

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

The tracking and association process is illustrated in Fig. 3. The output of Algorithm 1 is the $(k-1)$th superresolved frame, $\mathbf{i}_{k-1}$, whose nonzero values correspond to the positions of the MBs present in this frame. Next, given all previously known tracks $\mathbf{T}^1, \ldots, \mathbf{T}^{P_{k-1}}$, these positions need to be associated with 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 detected MB to one of the known P_{k-1} 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 [38], is considered one of the most popular data association techniques [28] and has been applied to a wide variety of multi-target tracking problems [39]. 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 performed by formulating and solving the maximum weighted independent set [40], [41], for example. The likelihood calculation relies on all past observations of each target [40]. 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 are acquired. As such, given the latest measurements in
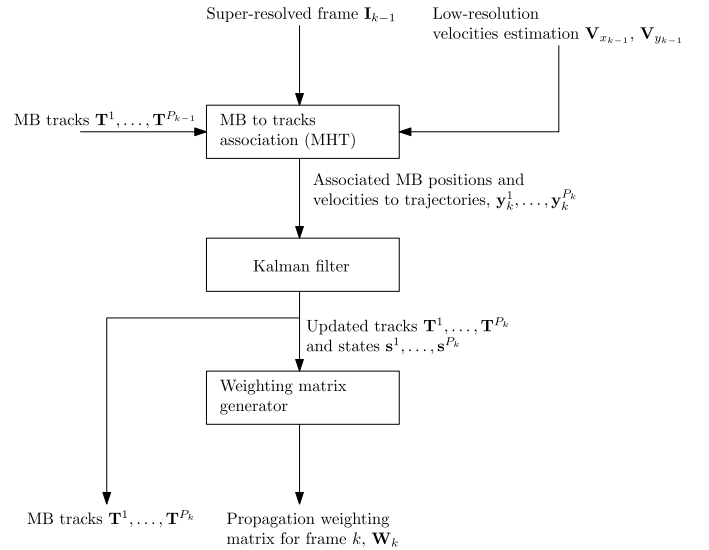

Fig. 3. Automatic tracking and data association procedure (detailed building blocks of the "automatic tracking algorithm" block in Fig. 2). New detected MBs from the $(k-1)$th superresolved frame are associated with previously known P_{k-1} tracks, or open new tracks, while nonassociated tracks are closed (uppermost block, realized by the MHT algorithm). Then, using the 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 OF on the low-resolution $k-1$ frame are associated with the newly detected MBs and used as measured velocities for the Kalman filter update. Thus, an updated track estimation is produced. Finally, the 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.

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). In practice, we use the Lisbon implementation, taken from [42], [43], which offers full integration into the MATLAB environment. Due to its flexibility, the integration of this implementation with the specific problem at hand, in this case, 2-D tracking of MBs, is easy and fast.

At the end of this association stage, existing tracks have been assigned new measurements (MBs positions and velocities), and new tracks are generated, if new MBs were detected. 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. The next stage of 3SAT is measurement integration with corresponding tracks, and propagation of the updated tracks to frame k.

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

$$\mathbf{s}_k^p = \mathbf{\Phi} \mathbf{s}_{k-1}^p + \boldsymbol{\eta}_k^p \tag{5}$$
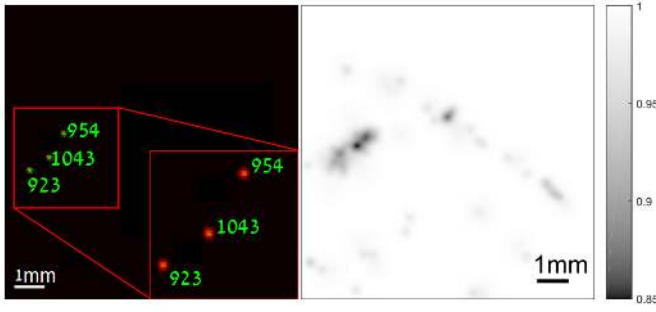
Fig. 4. Example of association of new measurements to existing tracks (left). In each frame, detected MBs are associated with track identification numbers that correspond either to existing or new tracks (MB to tracks association block in Fig. 3). Thus, the association of new data points to tracks occurs. Red box: enlarged area. MBs were smoothed slightly for visualization purposes only. An example of the (diagonal of the) weighting matrix $\mathbf{W}_k$ presented as an $N \times N$ image (weighting matrix generator block in Fig. 3) (right). Darker areas (lower values) correspond to more likely positions for MBs, and thus these pixels are given lower weights.

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* [44]. Ideally, a constant velocity model has zero acceleration, or zero second-order derivative. In practice, CWNA assumes that the velocity of each MB has slight perturbations, described by zero-mean white noise with power spectral density ρ. In (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 [44], the CWNA covariance matrix $\mathbf{Q}_k^p$ is given by

$$\mathbf{Q}_k^p = \begin{bmatrix} 1/3\Delta T^3 & q1/2\Delta T^2 & 0 & 0 \\ 1/2\Delta T^2 & \Delta T & 0 & 0 \\ 0 & 0 & 1/3\Delta T^3 & 1/2\Delta T^2 \\ 0 & 0 & 1/2\Delta T^2 & \Delta T \end{bmatrix} \rho$$

where ρ is chosen empirically.

The measurement model for the pth MB is then given by

$$\mathbf{y}_k^p = \mathbf{s}_{k-1}^p + \boldsymbol{\zeta}_k^p \qquad (6)$$

where $\boldsymbol{\zeta}_k^p$ 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, this matrix is chosen to be diagonal. As $\mathbf{R}_k^p$ is data dependent, the values of the diagonal are chosen empirically based on the expected velocity magnitudes of individual MBs in the data. As validation, these numbers are also compared with simulations of similar velocity magnitudes distribution, as exemplified in Section IV, Fig. 4.

From the superresolved image $\mathbf{i}_{k-1}$, we measure the position of the MBs. Specifically, consider an MB which is detected in position $[n_k \Delta_X, n_y \Delta_Y]$, where Δ_X and Δ_Y are the known sizes of each pixel in the superresolved 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 [23] on the low-resolution movie frames, as will be described in Section II-E.

The Kalman filter update rules are now formulated based on the propagation (5) and the measurement (6) models. MB state propagation to the next frame and its corresponding propagated estimation covariance matrix are given by

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

Using (7), the weighting matrix $\mathbf{W}_k$ is calculated as described in Section II-D. Next, (4) is minimized to recover the kth superresolved 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

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

and the innovation step along with the updated estimation error covariance matrix are

$$\begin{aligned} \mathbf{s}_{k|k}^p &= \mathbf{s}_{k|k-1}^p + \mathbf{K}_k^p \left(\mathbf{y}_k^p - \mathbf{s}_{k|k-1}^p \right) \\ \mathbf{P}_{k|k}^p &= \left(\mathbf{I}_{4\times4} - \mathbf{K}_k^p \right) \mathbf{P}_{k|k-1}^p. \end{aligned} \qquad (9)$$

From the innovation step (9), the states are updated as $\mathbf{s}_k^p = \mathbf{s}_{k|k}^p$ with estimation covariance matrix $\mathbf{P}_k^p = \mathbf{P}_{k|k}^p$.

D. Weighting Matrix Formulation

After the states for all MBs are propagated using (7) and associated with existing or new tracks, we turn to formulate the weighting matrix $\mathbf{W}_k$, as illustrated in Fig. 3 (bottom). The propagated state $\mathbf{s}_{k|k-1}^p$ represents the position and velocity of the pth MB and has its associated estimation error covariance matrix $\mathbf{P}_{k|k-1}^p$. Based on state predictions, a spatial MB-likelihood map $\mathbf{J}_k$ is formulated, 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_{k-1} 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

$$J_k[i,j] = \sum_{p=1}^{P_{k-1}} A^p e^{-q^p \left(\frac{1}{\sigma_x^{p2}} \left(i - x_0^p\right)^2 - c^p \left(i - x_0^p\right) \left(j - y_0^p\right) + \frac{1}{\sigma_y^{p2}} \left(j - y_0^p\right)^2 \right)} \qquad (10)$$

with $A^p = (|2\pi \mathbf{P}_{k|k-1}^p|)^{1/2}$, $[x_0^p, y_0^p] = [s_{k|k-1}^p[1], s_{k|k-1}^p[3]]$, $\sigma_x^p = P_{k|k-1}^p[1,1]$, $\sigma_y^p = P_{k|k-1}^p[3,3]$, $q^p = 1/(2(1-\rho^{p2}))$, $c^p = 2\rho^p/(\sigma_x^p \sigma_y^p)$ and $\rho^p = P_{k|k-1}^p[1,3]/(\sigma_x^p \sigma_y^p)$. The diagonal of the weighting matrix $\mathbf{W}_k$ is the inverse of the
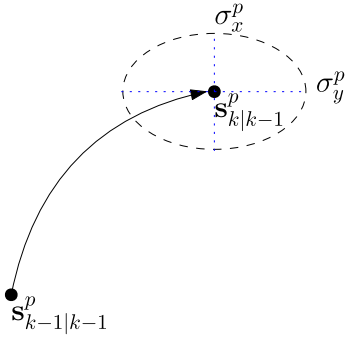
Fig. 5. Generation of the weighting matrix $\mathbf{W}_k$ [as described in (10) and (11)]. Previously estimated state $\mathbf{s}_{k-1|k-1}^p$ is propagated to state $\mathbf{s}_{k|k-1}^p$, according to (7). Its propagated error covariance matrix $\mathbf{P}_{k|k-1}^p$ is then used to draw an ellipse around its location, where $\sigma_x^p = \mathbf{P}_{k|k-1}^p[1, 1]$ and $\sigma_y^p = \mathbf{P}_{k|k-1}^p[3, 3]$. Aggregation of all propagated uncertainty ellipses generates an image of possible MBs locations. The matrix $\mathbf{W}_k$ is proportional to the inverse of this image.

Algorithm 2 3SAT

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

   Initialize $\mathbf{W}_1 = \mathbf{I}_{N^2 \times N^2}$

   Perform sparse superresolution on $\mathbf{z}_1$ using Algorithm 1

   for $k = 2, \ldots, K$ **do**

      Given $\mathbf{z}_k$ and $\mathbf{W}_{k-1}$, perform sparse super resolution on $\mathbf{z}_k$ using Algorithm 1 and estimate $\mathbf{i}_k$

      Estimate OF on $\mathbf{z}_k$ using MATLAB's opticalFlow command and estimate velocity components $\mathbf{V}_{x_k}$ and $\mathbf{V}_{y_k}$

      Given $\mathbf{i}_k$, $\mathbf{V}_{x_k}$ and $\mathbf{V}_{y_k}$, construct new measurement vectors for all P_k detected MBs $\mathbf{y}_k^p$, $p = 1, \ldots, P_k$

      Associate $\mathbf{y}_k^p$ to existing tracks $\mathbf{T}^p$, $p = 1, \ldots, P_k$/open new tracks/close old tracks using the MHT algorithm (Fig. 3)

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

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

      Given the updated tracks, construct weighting matrix $\mathbf{W}_k$ using (11)

   end for

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

---

vectorized form of $\mathbf{J}_k$ plus a regularization value ϵ, to avoid division by zero

$$W_k[i, i] = \frac{1}{J_k[\lfloor i/N \rfloor, (i \mod N)] + \epsilon}, \quad i = 1, \ldots, N^2$$

$$(11)$$

where $\lfloor \cdot \rfloor$ is the floor operation and $(x \mod y)$ is the modulo operation with the swap $0 \to N$. An illustration of such a weighting matrix can be observed in Fig. 4 (right). Vectorization of this $N \times N$ image is the diagonal of $\mathbf{W}_k$. The main building blocks of 3SAT are described in Algorithm 2.

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

E. Velocity Regularization Via Optical Flow Estimation

To improve the tracking procedure of individual MBs, velocity measurements are provided to the Kalman filter as part of the input to (6). This is done by OF estimation [23]–[26] 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 detected MB. This MB has a single position measurement. Without additional information on its general 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 available, then the filter will propagate the MB to a location in which the MB is more likely to be detected in the next frame.

OF estimation methods are easily implemented using the opticalFlow command in MATLAB. We achieved good performance with the method of Lukas and Kanade [26] 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 $N \times N$ superresolved images $\mathbf{I}_k$, and OF estimation is performed subsequently. This procedure ensures that each pixel in the superresolved 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 superresolved frame $\mathbf{I}_k$. Formally, the xy velocity fields obtained by OF estimation over the interpolated low-resolution frame $\mathbf{Z}_k$ are denoted 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 correspond to the pixel-wise lateral and axial estimated velocities, respectively. Next, for MBs detected in pixels $[i_p, j_p]$ from $\mathbf{I}_k$, we associate the corresponding velocity values from $\mathbf{V}_{x_k}$ and $\mathbf{V}_{y_k}$

$$y_k^p[2] = V_{x_k}[i_p, j_p]$$
$$y_k^p[4] = V_{y_k}[i_p, j_p] \quad p = 1, \ldots, P_k. \quad (12)$$

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

Note that OF estimation is performed on the low-resolution movie and not on the superresolved frames, as the basic assumption of OF, known as the pixel intensity consistency assumption [24] does not hold on the superresolved images. This is because a typical superresolved image looks like the image displayed in Fig. 4 (left). The enlarged box shows the localization of three MBs (smoothed only for display purposes). Typically, in the next frame, these MBs move considerably, which prevents reliable OF estimation. In contrast, a much more reliable OF estimation is achieved on the low-resolution images due to the spreading of the echoes from the MBs over several adjacent pixels.
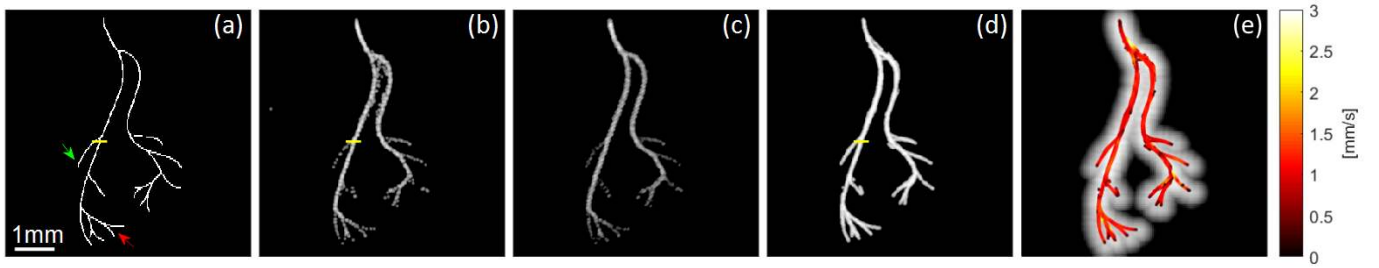
Fig. 6.   Simulation results. (a) Groundtruth image of bifurcating blood vessels. (b) Superlocalization reconstruction. (c) Superresolution sparse recovery obtained by minimizing (3) via FISTA. (d) 3SAT recovery by accumulating all recovered MB trajectories. (e) Superimposed velocity trajectories over the MIP image obtained from the 3SAT recovery. Yellow lines: selected profiles are presented in Fig. 7. All reconstructions are displayed in logarithmic scale with a dynamic range of 40 dB.

Since the velocities are measured on the low-resolution images, from non-resolved MBs, they do not represent the velocities of individually resolved MBs. Instead, they constitute a coarse, low-resolution estimate of the average velocities from the nonresolved MBs. As such, the velocity measurements are weighted with ten times larger values in the covariance matrix $\mathbf{R}_k^p$ than the position measurements.

III. MATERIALS AND METHODS

A. Numerical Simulations

Using MATLAB, we simulate a bolus injection of 140 MBs into vascular bifurcations over an acquisition period of 96 frames, with the frame rate of 10 Hz (total acquisition time is 10 s), and pixel size of 0.15×0.15 mm^2. Acquisition of the MBs was simulated by summing RF-modulated PSFs for each of the MB locations. RF lines were subsequently demodulated using the Hilbert transform. Images were formed by subsequentially taking the absolute value of the demodulated data. The received modulation frequency is 7 MHz (second harmonic of 3.5 MHz, similar to our *in vivo* acquisition setup), with a Gaussian PSF having a standard deviation of 0.14 mm in the axial direction and 0.16 mm in the lateral direction. The entire acquisition process was approximated by a linear mapping from MB locations to beamformed image, through the PSF. The extracted second harmonic component was simulated through the use of an RF modulation at that frequency. MBs flow from the highest point of Fig. 6(a) down to the terminals of all branches. To simulate the bolus injection, the probability for an MB to appear at the point of origin follows a Gaussian distribution in the following manner. MB velocities' magnitudes and directions were generated by taking the maximum between values drawn from a normal distribution with a mean of 1 mm/s and standard deviation of 1 mm/s, and zero, to avoid negative magnitude values. Additive white Gaussian noise with a standard deviation of 0.03 was added.

We set $P = 8$ and recover the superresolved images on an eight times denser grid than the low-resolution grid. The regularization parameter was chosen to be $\lambda = 0.002$. We iterate over 4000 iterations per frame and set $\epsilon = 1$, $\rho = 260$, and $\mathbf{R}_k = \text{diag}\{0.1, 1, 0.1, 1\}$. In the MHT algorithm, the probability for not detecting an existing target was chosen as 0.1, the probability for a new target to appear is 0.2, and

the probability for false alarm is 0.01. A maximum number of six leaves are used.

B. In Vivo Experiments

The CEUS data of a human prostate from a patient who underwent radical prostatectomy were acquired at the AMC University Hospital (Amsterdam, the Netherlands), using a 2-D transrectal US probe (C10-3v) and an iU22 scanner (Phillips Healthcare, Bothell, WA, USA). The scanner operated in a contrast-specific mode at a frame rate of $1/\Delta T = 10$ Hz with central transmission frequency of 3.5- and 7-MHz central frequency upon reception of the second harmonic. A 2.4-mL MB bolus of SonoVue (Bracco, Milan, Italy) was administered intravenously, and 100 frames (10 s) were collected for further analysis. The pixel size is 0.146×0.146 mm^2 and the mechanical index was set to 0.06 to attain sufficient SNR, while limiting MB destruction [45], [46]. The study was approved by the local ethics committee of the Academical Medical Center, University Hospital of Amsterdam, the Netherlands. All patients signed informed consent.

We consider two examples taken from the *in vivo* scan. For all experiments, we set $P = 4$ and recover the superresolved images on four times denser grid than the low-resolution grid. Since variations in PSF shape and orientation are inevitable over large fields of view, as in the case when scanning with a transrectal probe, for each case, the PSF was estimated separately. 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, the probability for not detecting an existing target was set to 0.1, the probability for a new target to appear to 0.5, the probability for false alarm to 0.01 and a maximum number of leaves to 6. All superlocalization reconstructions were performed after applying wavelet-based denoising to suppress residual tissue signal, as the power modulation scheme does not remove the tissue signal completely. Without this denoising step, many false positive detections were observed. FISTA and 3SAT recoveries were performed on the separated CEUS signal, without any prior denoising.

IV. RESULTS

A. Numerical Simulations

Fig. 6 shows the reconstruction results of the simulated data set of flowing MBs within a simulated vascular net-
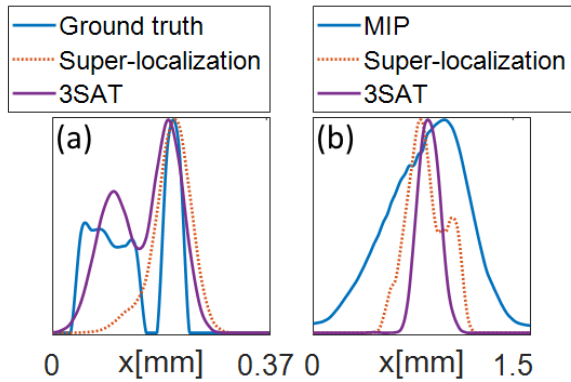
Fig. 7. (a) Intensity profiles (a.u.) taken along the yellow lines in Fig. 6. (b) Intensity profiles (a.u.) taken along the yellow lines in Fig. 10.

TABLE I
PEAK-TO-PEAK DISTANCES FOR THE INTENSITY PROFILES IN FIG. 6 (FOR 3SAT, WITHOUT/WITH POSTPROCESSING SMOOTHING)

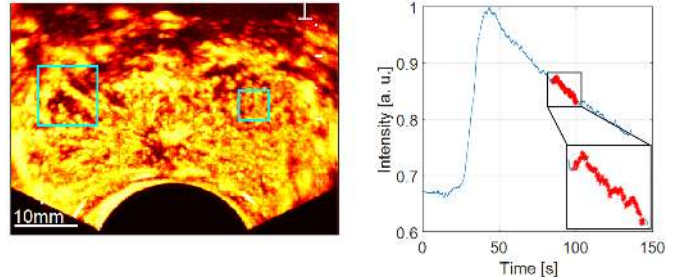|	ground truth	3SAT	Super-localization
Solid line [μm]	71.5	82.8/95	—

Fig. 8. MIP image of the *in vivo* prostate scan used in this study (left). Right rectangle: first examined the area of the prostate (Fig. 9). Left rectangle: second area (Fig. 10). Time–intensity profile calculated as the mean frame intensity as a function of scan time (right). Red highlighted region: processed time segment in both examples. This area exhibits recirculation of MBs in the patient's bloodstream. Time–intensity curve for the right rectangle is shown in Fig. 12(d), while the corresponding curve for the left rectangle is shown in (e).

work. Fig. 6(a) shows the groundtruth architecture, while Fig. 6(b)–(d) shows the reconstruction results of superlocalization [47], sparsity-driven superresolution [20] [minimizing (3)] and 3SAT, respectively. Fig. 6(e) shows an overlay of MB trajectories, colored by their estimated velocities over the maximum intensity projection (MIP) image.

Visual inspection reveals that 3SAT recovery [Fig. 6(d)] seems the smoothest and most continuous, that is depicting a more complete image of the underlying vascular network, compared with the superlocalization (which seems very discontinuous and includes false positive detections) and sparsity-based reconstructions, depicted in Fig. 6(b) and (c), respectively. The green arrow in Fig. 6(a) indicates a bifurcating blood vessel, which is almost nondepicted in the superlocalization image Fig. 6(b) and is discontinuous in Fig. 6(c) of the sparsity-driven approach. Conversely, 3SAT Fig. 6(d) detects this blood vessel completely, showing a continuous connection to the main blood vessel. The red arrow indicates another example of clear bifurcation depictions by 3SAT, which are discontinuous in the other reconstructions.

Fig. 6(e) shows that the estimated velocities are in the range 0–3 mm/s. A histogram of the measured velocities is shown in panel (c) of Fig. 11, where the velocities distribution is indeed between 0 and 3 mm/s. This histogram is compared with the scaled (truncated) Gaussian distribution from which MB velocities where generated (solid transparent curve). A good match between the two distributions is achieved, validating the performance and reliability of 3SAT.

In panel (a) of Fig. 7, selected intensity profiles (a.u.) were measured along the dashed yellow lines in Fig. 6(a), (b), and (d). In this example, it is evident that the 3SAT profile detects the two peaks (vessel branching) also present in the ground truth, with good agreement, while the superlocalization procedure fails to detect the rightmost peak. This situation is expected since most frames consist of overlapping MBs, especially at bifurcations. In scenarios of extensive overlap, superlocalization tends to result in aggregations of detections, often leading to mis-detections of nearby vessels, as opposed to sparsity-based techniques that are designed to account for this overlap. Table I gives the comparison of the peak-to-peak distance for the bifurcations

along the yellow lines in Fig. 6, as measured in the ground truth image, superlocalization, and 3SAT recoveries. Imaging at 7 MHz, the imaging wavelength λ_0 is 220 μm (speed of sound 1540 m/s). A common measure of resolution is the full-width-at-half-max (FWHM), which for a Gaussian function is approximately 2.355σ, or in case of the simulation 330 μm in the lateral direction. Thus, the values attained by 3SAT, in this example, correspond to a separation distance smaller than $\lambda_0/2$, and 3.5–4 times improvement in the FWHM, demonstrating that 3SAT achieves subwavelength imaging.

B. In Vivo Experiments

In this section, we present *in vivo* reconstruction results of 3SAT. US acquisition parameters and reconstruction parameters are given in Section III. Fig. 8 (left) shows the MIP image of the entire prostate, while the rectangles correspond to the selected processed areas. In this study, we focus on selected patches and do not process the entire aperture of the prostate, since we use a curved transrectal probe which spans a wide field of view (FOV). Since the FOV is large, the PSF depends on the scanned region (e.g., the PSF for the right rectangle is not the same for the left one). A main assumption we make is that of a linear, shift-invariant acquisition model, hence we process smaller patches over which this assumption seems to hold. For each processed area the PSF is estimated from the data, independently. The right figure shows the time–intensity curve calculated as the mean intensity at each frame from the entire FOV. The red area marks the observation window over which processing is performed. This time span was chosen during the wash-out phase of the MBs but also exhibits recirculation of MBs and an increase in MB concentration at the beginning of the acquired period.
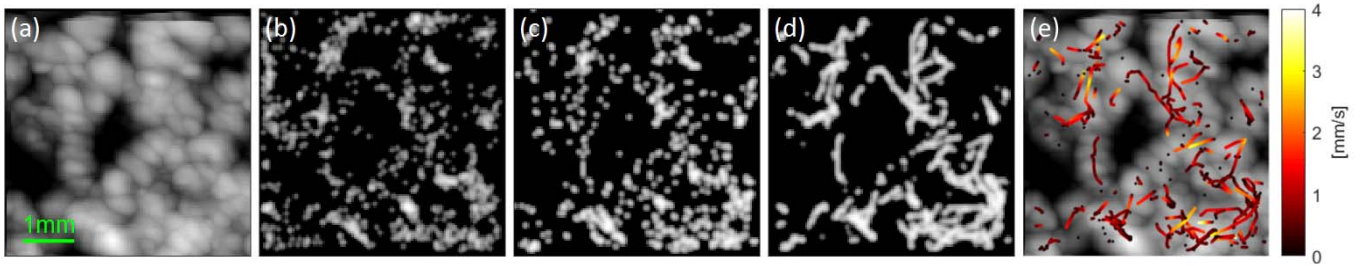
Fig. 9.   3SAT applied to an *in vivo* scan from a human prostate. (a) MIP image from 100 frames. (b) Superlocalization recovery. (c) Superresolution sparse recovery obtained by minimizing (3) via FISTA. (d) 3SAT recovery by accumulating all recovered MB trajectories. (e) Superimposed velocity trajectories over the MIP image obtained from the 3SAT recovery. All images are displayed in logarithmic scale with a dynamic range of 30 dB.
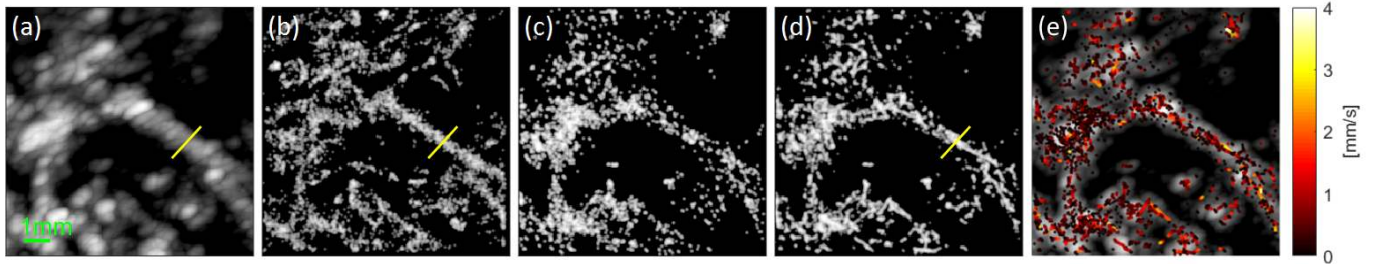

Fig. 10.   Additional example of 3SAT recovery of an *in vivo* human prostate scan. (a) MIP image from 100 frames. (b) Superlocalization recovery. (c) Superresolution sparse recovery obtained by minimizing (3) via FISTA. (d) 3SAT recovery by accumulating all recovered MB trajectories. (e) Superimposed velocity trajectories over the MIP image obtained from the 3SAT recovery. All images are displayed in logarithmic scale with a dynamic range of 30 dB.

Figs. 9 [right rectangle in Fig. 8 (left)] and 10 [left rectangle in Fig. 8 (left)] compare between different reconstructions in two areas of a prostate CEUS scan. In both Figs. 9 and 10, (a) shows the MIP image. This image is diffraction limited and was generated as reference for standard non superresolution image processing by taking the pixelwise maximum value over the entire movie. Fig. 9(b) shows the resulting superlocalization-based image, by localizing individual MBs per frame. Fig. 9(c) depicts sparsity-based superresolution as obtained by minimizing (3) via FISTA, while Fig. 9(d) shows the 3SAT output. Finally, Fig. 9(e) displays an overlay of the estimated velocities' trajectories on the MIP image.

We first consider Fig. 9. By qualitative consideration, the comparison of (b)–(d) and (a) shows that all methods seem to achieve superresolution, with a relatively good agreement among them. That is, main superresolved features seem to be present in all three methods, such as a small vertical blood vessel located at the bottom left portion of the images of (b)–(d). However, the 3SAT image (d) qualitatively appears smoother and more continuous, showing distinct trajectories that are absent in the sparse recovery and superlocalization images. These consist of a larger set of disconnected and isolated MBs. Moreover, Fig. 9(e) presents the velocity magnitude estimations from the tracked MBs. The vast majority of obtained MB flow velocities are on the order of up to 1–2 mm/s, in line with the previous observations on blood flow in microvessels [15]. This is also confirmed by the velocity magnitude histograms, displayed in Fig. 11(a). However, additional higher velocities up to ~4 mm/s are also observed. Such velocities may correspond to larger vessels, with increased flow.
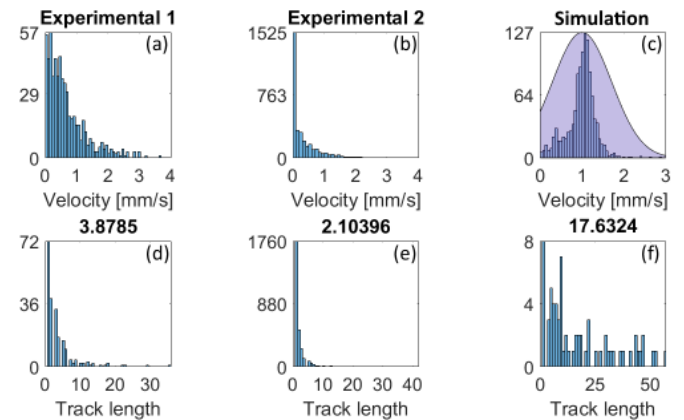

Fig. 11.   Top row: estimated velocity histograms. (a) Experimental data set shown in Fig. 9. (b) Experimental data set shown in Fig. 10. (c) Simulation data set shown in Fig. 6. Transparent solid curve indicates a scaled Gaussian distribution with mean and standard deviation of 1. Bottom row: track length histogram of corresponding data sets. (d) Experimental data set of (a). (e) Experimental data set of (b). (f) Simulated data set of (c). Track lengths are measured as the number of measurements associated with each track. Numbers correspond to the mean track length of each histogram.

Considering Fig. 10(b)–(d), qualitatively show once more superresolution imaging of the prostate vasculature, compared with the MIP image of (a). These figures further support the conclusions drawn in Fig. 9. In this example, we also quantify the resolution increase of superlocalization and 3SAT, compared with the MIP image. The yellow lines indicate a selected intensity profile of a blood vessel. This vessel can clearly be seen in the MIP image. Fig. 7(b) shows the
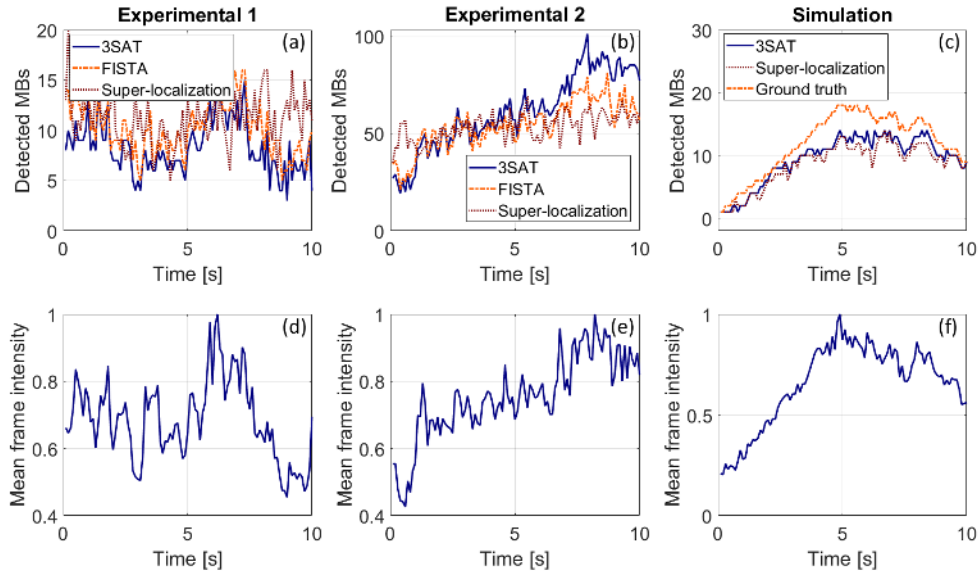
Fig. 12. Detected MBs per frame for both experimental data sets shown in (a) Fig. 9, (b) Fig. 10, and (c) simulation. (d)–(f) Corresponding time–intensity curves for each scan. These curves were calculated as the framewise mean pixel intensity over time.

corresponding intensity profiles of the MIP, superlocalization, and 3SAT recoveries (arbitrary units, intensity normalized to one). Both superlocalization and 3SAT achieve narrowing of the blood vessel contour, as compared with the MIP profile (blue solid line).

Fig. 10(e) depicts pointlike and short trajectories, with low velocities, alongside longer and smoother trajectories. Such pointlike trajectories are attributed to the fact that the vascular bed of the prostate is inherently 3-D, 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. Further quantification is also shown in Fig. 11(d)–(f), showing a histogram of track lengths (bold numbers of above these figures indicate mean velocity length), in which larger track lengths are observed in the simulation, and shorter lengths in the experimental data sets.

We further quantify the number of detected MBs in each frame, for each of the superresolution methods discussed above. Fig. 12(a)–(c) indicates the number of detected MBs per frame for the experimental data presented in Figs. 9 and 10 and the simulation shown in Fig. 6. The lines indicate the number of detected MBs for 3SAT (solid blue), sparsity-based superresolution via FISTA [dash-dot orange, (a) and (b)] and superlocalization (dot brown). The dash-dot orange line in (c) indicates the ground truth number of MBs in the simulation.

Considering (a), it is evident that all three methods perform similarly, and detect a similar number of MBs in each frame. In this case, MB density is such that all methods perform similarly. On the other hand, (b) shows an increasing number of detected MBs for all the methods. After ∼35 frames, 3SAT clearly detects more MBs compared with superlocalization and FISTA-based sparse recovery (first ∼10 frames show higher detection rate for superlocalization, but this is likely the result of false detections due to residual tissue signal or noise in the

CEUS data). This scan was taken during the wash-out phase of MBs circulation; however, the increase in detections indicates that this data is a part a recirculation of the MBs, in which MB concentration increases as time progresses. The simulation corresponding to (c) simulates the entire circulation of MBs in the blood steam, from the wash-in phase up to the wash-out phase. Fig. 12(c) shows that 3SAT is able to consistently detect more MBs, achieving an improved estimation of the number of MBs per frame, as compared with the ground truth curve. The discrepancy between the ground truth and 3SAT curves is due to aggregated MBs, which cannot be resolved by any of the methods. In (c), corresponding detections of the FISTA method are not displayed due to a high rate of false detections.

Fig. 12(d)–(f) shows the corresponding time–intensity curves for each scan, measured as the framewise mean pixel intensity over time. In [48], a linear relation between the intensity and MB concentration was verified for concentrations up to 1 mg/L. Visual inspection shows that the time–intensity curves in the bottom figure correlate with the 3SAT curves in the corresponding (a)–(c), further supporting the conclusion that 3SAT accurately detects the number of MBs per frame, as compared with superlocalization, for such concentrations. Most notable is the curve in (e), which matches that of 3SAT, showing an increase in MB detections over time. Furthermore, it shows that indeed the high number of detections by the superlocalization method in the first ∼10 frames is most likely the result of false detections.

All panels support the conclusion that at low MB concentrations (∼2–3 MB/mm^2), 3SAT performs similar to superlocalization, but exhibits improved performance, when the density increases (∼10 MB/mm^2), which is about $3-5$ times denser MB concentration.[2]

---

[2]Concentration values were measured by dividing the number of detected MBs per frame with each respective FOV.

V. DISCUSSION AND CONCLUSION

This is the first work to exploit the inherent motion kinematics of individual MBs as a structural prior for superresolution. Since individual MBs flow within blood vessels, their positions can be predicted from one frame to the next. 3SAT exploits this additional information to improve sparse recovery, by solving a support aware minimization problem, as formulated in (4). Using Kalman filtering, 3SAT is able to track and propagate the trajectories of individual MBs from one frame to the next. Moreover, we introduce velocity measurements via OF estimation to improve the tracking process for superresolution imaging. In a recent study, de Senneville *et al.* [49] introduced a similar concept of MB transport to quantify the velocity amplitude of bolus arrival in CEUS as a diagnostic tool. 3SAT extends these ideas by relying on OF as measurements, while estimating individually resolved MBs' velocities. Figs. 9 and 10 show the power and potential of 3SAT on *in vivo* data. Both clear and smooth superresolution imaging are achieved, as well as a quantitative measurement of the flow velocities of individual MBs, quantification of resolution enhancement and MB detection rates. This study shows that the 3SAT approach achieves a higher and more accurate detection rate of MBs than the superlocalization-based approach, when MB density increases, as depicted in Fig. 12.

3SAT operates well with high MB concentrations (e.g., ∼10 MB/mm^2), for which significant MB overlap is present, as quantified in Fig. 12. However, the detection rate of MBs is noise dependent. In very noisy scenarios, 3SAT is not guaranteed to recover all individual MBs, although in the experiments showed in this paper, detection rates are high. By exploiting the sparse nature of the individual MB echoes, 3SAT is able to depict the vasculature with a relatively low number of frames. In our *in vivo* experiments we used two data sets of 100 frames. However, as MB density increases even further, several mechanisms of 3SAT may 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. Moreover, high MB velocities relative to the scanner frame rate may also limit the correct association of MBs to tracks. Finally, 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 reported in this work, for clinical bolus doses and 10-Hz scanners, 3SAT depicts a smoother and more consistent vasculature and is able to detect more MBs than the state-of-the-art techniques in high concentration scenarios.

As can be observed in Fig. 8 (right), the processed acquisition period was taken during the wash-out phase of the MBs.

This period was chosen, and correspondingly the two subregions, in order to test the performance of 3SAT in relatively low and constant densities (Fig. 9) and in higher densities (Fig. 10). In Fig. 10, recirculation of MBs appears, showing an increase in MB signal and density. These two subregions serve to assess 3SAT in both scenarios and compare its performance to superlocalization. In fact, Fig. 12 demonstrates both *in vivo* and *in silico* that as MB density increases, 3SAT is able to reliably recover more MBs than superlocalization. Furthermore, since the removal of tissue signal is not perfect, it seems that superlocalization is more prone to false detections than sparsity-based approaches. This happens since sparsity-based approaches use explicit information about the measured PSF, whereas superlocalization does not.

Velocity estimates of 3SAT may vary when considering different periods during the entire MBs circulation period. This may happen as MB velocities may differ during the wash-in and wash-out phase. Furthermore, considering longer acquisition periods may reveal finer blood vessels as statistically, the likelihood of MBs to flow through these vessels increases. However, such longer durations may lead to poorer MB detection and velocity estimation, as motion becomes more dominant and the registration process might fail. Minimization of these errors further motivates the use of high MB concentrations and the development of methods which can compensate for MBs overlap, such as 3SAT.

In [19], a similar model to (1) was introduced over the beam-formed complex in-phase and quadrature signal (in which acquired echoes from different MBs are superimposed together), while this work assumes such a model over the real-valued intensity images. Although intensity images are formed through a nonlinear operation of envelope detection, in practice, both in simulations and *in vivo* experiments, we observed good reconstruction performance using model (1), as 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 3SAT. Second, it is known that capillary flow is nonturbulent [8], [22]. Thus, a simple linear propagation model is reasonable for the tracking procedure.

There are several limitations to 3SAT. The first is inherent to all US superresolution techniques. MBs must flow through the vasculature in order to detect it, thus setting a minimal acquisition time for any superresolution imaging technique. In 3SAT, 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, 3SAT includes several parameters which should be selected properly, among which are the sparsity regularization parameters λ, ϵ, ρ and the probabilities for the MHT algorithm. In this work, these parameters were chosen manually, but according to some important guidelines. Typically ϵ should be relatively small, similar to the values we chose, to avoid numerical errors, but ultimately may vary from one scan to the other. Next, the second and fourth entries of the diagonal

of $\mathbf{Q}_k^p$ correspond to the variance of the velocity estimates in the x- and y-directions, respectively. Therefore, in fact, the standard deviation of each velocity estimate is proportional to $(\Delta T \rho)^{1/2}$. This suggests guidelines to the determination of the power spectral density of the noise, ρ. In our *in vivo* examples and based on past literature, see [22], we know that blood flow in small vessels is typically 0–2 mm/s and can be higher (e.g., 4 mm/s as we measure in some parts of our *in vivo* examples) in larger vessels. Thus, ρ is set according to the expected velocities in the imaged medium. Since $\Delta T = 0.1$ s, choosing $\rho = 500$ results in a standard deviation of 7 mm/s, which accommodates for slight variations in the estimated velocity values of our model. The probabilities for the MHT algorithm were chosen to be similar to those specified in the MATLAB example provided with the code package.

The selection of the parameter λ corresponds to the given SNR and the number of expected MBs in the image. This number can of course change between the wash-in and wash-out phases of MBs flow, as MB concentration changes. In our experiments, this parameter was chosen according to an upper bound on the expected number of MB per mm^2. However, this number can also change between scans of different organs or different parts of the same organ, and in this case, additional validation and testing will have to be performed.

As was also mentioned in [8], 3SAT is designed to work with 2-D images but is affected by the inherent 3-D geometry of blood vessels going in and out of the imaging plane. Thus, some of the detected MBs cannot be tracked over several consecutive frames, resulting in single detections, as was presented in Figs. 9 and 10. This is an inherent limitation to all 2-D-based superresolution techniques. It can probably be alleviated with the introduction of 3-D probes and volumetric scans, which will enable 3-D tracking of individual MBs.

Before concluding, we would like to discuss some computational aspects of online sparse tracking, as the number of detected MBs grows. Angelosante *et al.* [32] 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 3SAT applies Kalman filtering to each detected 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), so that 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. Ackermann and Schmitz [8] considered a modified version of Markov Chain Monte Carlo (MCMC) data association [50] to account for this growth. The computational complexity of MHT can also be controlled by limiting the pruning depth, achieving a tradeoff between accurate data association and computational complexity. In general, any automatic association algorithm may be used in the uppermost block in Fig. 3 instead of MHT, such as the joint probabilistic data association (JPDA) [51] or the MCMC algorithm of [8].

To conclude, in this work, we presented a new algorithm to improve sparsity-based superresolution CEUS imaging, taken mainly from low-frame rate clinical scanners. By formulating a weighted sparse recovery minimization problem, combined with online tracking of individual MBs, we are able to improve the sparse recovery process and fill-in for additional information of MB positions. 3SAT achieves a smoother depiction of the vasculature and provides quantitative information regarding MB kinematics and MB detection rates. We applied our algorithm to both simulations and *in vivo* human prostate scans, obtained from low-frame rate (10 Hz), clinically approved US machines, demonstrating superresolution recovery of the vascular bed with 100 frames. On these scans, we demonstrate that as MB density increases, 3SAT is able to recover more MBs compared with the state-of-the-art methods, while achieving superresolution imaging. Since 3SAT employs an online estimation process, it may be suitable for real-time applications within commercially available US machines.

REFERENCES

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

[2] N. de Jong, F. T. Cate, C. Lancée, J. R. T. C. 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," *Eur. J. Radiol.*, vol. 60, no. 3, pp. 324–330, Dec. 2006.

[4] J. M. Hudson *et al.*, "Dynamic contrast enhanced ultrasound for therapy monitoring," *Eur. J. Radiol.*, vol. 84, no. 9, pp. 1650–1657, Sep. 2015.

[5] C. Errico *et al.*, "Ultrafast ultrasound localization microscopy for deep super-resolution vascular imaging," *Nature*, vol. 527, no. 7579, pp. 499–502, Nov. 2015.

[6] M. A. O'Reilly and K. Hynynen, "A super-resolution ultrasound method for brain vascular mapping," *Med. Phys.*, vol. 40, no. 11, 2013, Art. no. 110701.

[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 Trans. Med. Imag.*, vol. 34, no. 2, pp. 433–440, Feb. 2015.

[8] D. Ackermann, G. Schmitz, and S. Member, "Detection and tracking of multiple microbubbles in ultrasound B-mode images," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 63, no. 1, pp. 72–82, Jan. 2016.

[9] T. Opacic *et al.*, "Motion model ultrasound localization microscopy for preclinical and clinical multiparametric tumor characterization," *Nature Commun.*, vol. 9, no. 1, 2018, Art. no. 1527.

[10] 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," *Sci. Rep.*, vol. 7, no. 1, 2017, Art. no. 13662.

[11] 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 Trans. Med. Imag.*, vol. 36, no. 1, pp. 169–180, Jan. 2017.

[12] F. Lin, S. E. Shelton, D. Espíndola, J. D. Rojas, G. Pinton, and P. A. Dayton, "3-D ultrasound localization microscopy for identifying microvascular morphology features of tumor angiogenesis at a resolution beyond the diffraction limit of conventional ultrasound," *Theranostics*, vol. 7, no. 1, p. 196, 2017.

[13] E. Betzig *et al.*, "Imaging intracellular fluorescent proteins at nanometer resolution," *Science*, vol. 313, no. 5793, pp. 1642–1645, 2006.

[14] 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.

[15] R. J. G. 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," *Med. Image Anal.*, vol. 35, pp. 610–619, Jan. 2017.

[16] Y. C. Eldar, *Sampling Theory: Beyond Bandlimited Systems*. Cambridge, U.K.: Cambridge Univ. Press, 2015.

[17] A. Bar-Zion, O. Solomon, C. Tremblay-Darveau, D. Adam, and Y. C. Eldar, "Sparsity-based ultrasound super-resolution imaging," in *Proc. 23rd Eur. Symp. Ultrasound Contrast Imag.*, 2017, pp. 156–157.

[18] O. Solomon, A. Bar-Zion, D. Adam, and Y. C. Eldar, "Fast and background free super-resolution ultrasound angiography," in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, Sep. 2017, p. 1.

[19] A. Bar-Zion, O. Solomon, C. Tremblay-Darveau, D. Adam, and Y. C. Eldar, "SUSHI: Sparsity-based ultrasound super-resolution hemo-dynamic imaging," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 65, no. 12, pp. 2365–2380, Dec. 2018.

[20] R. J. van Sloun, O. Solomon, Y. C. Eldar, H. Wijkstra, and M. Mischi, "Sparsity-driven super-resolution in clinical contrast-enhanced ultra-sound," in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, Sep. 2017, pp. 1–4.

[21] R. J. G. 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 Trans. Med. Imag.*, vol. 36, no. 3, pp. 826–837, Mar. 2017.

[22] J. A. Jensen, *Estimation of Blood Velocities Using Ultrasound: A Signal Processing Approach*. Cambridge, U.K.: Cambridge Univ. Press, 1996.

[23] J. L. Barron, D. J. Fleet, and S. S. Beauchemin, "Performance of optical flow techniques," *Int. J. Comput. Vis.*, vol. 12, no. 1, pp. 43–77, 1994.

[24] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artif. Intell.*, vol. 17, nos. 1–3, pp. 185–203, Aug. 1981.

[25] S. Baker and I. Matthews, "Lucas-Kanade 20 years on: A unifying framework," *Int. J. Comput. Vis.*, vol. 56, no. 3, pp. 221–255, 2004.

[26] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proc. Int. Joint Conf. Artif. Intell.*, 1981, pp. 674–679.

[27] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. ASME, D, J. Basic Eng.*, vol. 82, pp. 35–45, 1960.

[28] Y. Bar-Shalom and X.-R. Li, *Multitarget-multisensor Tracking: Princi-ples and Techniques*, vol. 19. London, U.K.: YBS, 1995.

[29] V. Mor-Avi, E. G. Caiani, K. A. Collins, C. E. Korcarz, J. E. Bednarz, and R. M. Lang, "Combined assessment of myocardial perfusion and regional left ventricular function by analysis of contrast-enhanced power modulation images," *Circulation*, vol. 104, no. 3, pp. 352–357, 2001.

[30] N. Vaswani and W. Lu, "Modified-CS: Modifying compressive sens-ing for problems with partially known support," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4595–4607, Sep. 2010.

[31] N. Vaswani, "Kalman filtered compressed sensing," in *Proc. 15th IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2008, pp. 893–896.

[32] D. Angelosante, S. I. Roumeliotis, and G. B. Giannakis, "Lasso-Kalman smoother for tracking sparse signals," in *Proc. Conf. Rec. 43rd Asilomar Conf. Signals, Syst. Comput.*, 2009, pp. 181–185.

[33] D. Angelosante, G. B. Giannakis, and E. Grossi, "Compressed sensing of time-varying signals," in *Proc. 16th Int. Conf. Digit. Signal Process.*, 2009, pp. 1–8.

[34] O. Solomon, M. Mutzafi, M. Segev, and Y. C. Eldar, "Sparsity-based super-resolution microscopy from correlation information," *Opt. Express*, vol. 26, no. 14, pp. 18238–18269, 2018.

[35] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imag. Sci.*, vol. 2, no. 1, pp. 183–202, 2009.

[36] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Appli-cations*. Cambridge, U.K.: Cambridge Univ. Press, 2012.

[37] E. J. Candès, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *J. Fourier Anal. Appl.*, vol. 14, nos. 5–6, pp. 877–905, 2008.

[38] D. B. Reid, "An algorithm for tracking multiple targets," *IEEE Trans. Autom. Control*, vol. AC-24, no. 6, pp. 843–854, Dec. 1979.

[39] G. Cybenko, V. Crespi, and G. Jiang, "What is trackable?" *Proc. SPIE*, vol. 6201, May 2006, Art. no. 620107.

[40] C. Kim, F. Li, A. Ciptadi, and J. M. Rehg, "Multiple hypothesis tracking revisited," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 4696–4704.

[41] P. M. Pardalos and N. Desai, "An algorithm for finding a maximum weighted independent set in an arbitrary graph," *Int. J. Comput. Math.*, vol. 38, nos. 3–4, pp. 163–175, 1991.

[42] D. M. Antunes, D. M. de Matos, and J. Gaspar, "A library for implement-ing the multiple hypothesis tracking algorithm," 2011, *arXiv:1106.2263*. [Online]. Available: https://arxiv.org/abs/1106.2263

[43] D. M. Antunes, D. Figueira, D. M. Matos, A. Bernardino, and J. Gaspar, "Multiple hypothesis tracking in camera networks," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops (ICCV Workshops)*, Nov. 2011, pp. 367–374.

[44] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Appli-cations to Tracking and Navigation: Theory Algorithms and Software*. Hoboken, NJ, USA: Wiley, 2004.

[45] N. de Jong, P. J. A. Frinking, A. Bouakaz, and F. J. T. Cate, "Detection procedures of ultrasound contrast agents," *Ultrasonics*, vol. 38, no. 1, pp. 87–92, Mar. 2000.

[46] S. B. Feinstein, "The powerful microbubble: From bench to bed-side, from intravascular indicator to therapeutic delivery system, and beyond," *Amer. J. Physiol.-Heart Circulatory Physiol.*, vol. 287, no. 2, pp. H450–H457, 2004.

[47] M. Ovesný, 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.

[48] M. P. J. Kuenen, M. Mischi, and H. Wijkstra, "Contrast-ultrasound diffusion imaging for localization of prostate cancer," *IEEE Trans. Med. Imag.*, vol. 30, no. 8, pp. 1493–1502, Aug. 2011.

[49] B. D. de Senneville *et al.*, "Development of a fluid dynamic model for quantitative contrast-enhanced ultrasound imaging," *IEEE Trans. Med. Imag.*, vol. 37, no. 2, pp. 372–383, Feb. 2018.

[50] S. Oh, S. Russell, and S. Sastry, "Markov chain Monte Carlo data association for multi-target tracking," *IEEE Trans. Autom. Control*, vol. 54, no. 3, pp. 481–497, Mar. 2009.

[51] T. E. Fortmann, Y. Bar-Shalom, and M. Scheffe, "Multi-target track-ing using joint probabilistic data association," in *Proc. 19th IEEE Conf. Decis. Control Including Symp. Adapt. Process.*, Dec. 1980, pp. 807–812.