Efficient Tracking of Sparse Signals via an Earth Mover’s Distance Dynamics Regularizer

Nicholas P. Bertrand*, Adam S. Charles*, Member, IEEE, John Lee*, Pavel B. Dunn, Christopher J. Rozell Senior Member, IEEE

Abstract—Sparse signal models have enjoyed great success, achieving state-of-the-art performance in many applications. Some algorithms further improve performance by taking advantage of temporal dynamics for streaming observations. However, the tracking regularizers are often based on the $\ell_p$-norm which does not take full advantage of the relationship between neighboring signal elements. In this work, we propose the use of the earth mover’s distance (EMD) as an alternative tracking regularizer for causal tracking when there is a natural geometry to the coefficient space that should be respected (e.g., meaningful ordering). Our proposed earth mover’s distance dynamic filtering (EMD-DF) algorithm is a causal approach to tracking time-varying sparse signals that includes two variants: one which uses the traditional EMD as a tracking regularizer for sparse nonnegative signals, and a relaxation which allows for complex valued signals. In addition, we present a computationally efficient formulation of EMD-DF (based on optimal transport theory), improving computational scalability for reasonably sized image-tracking applications. Through a series of simulations, we demonstrate the advantages of EMD-DF compared to existing methods on tracking sparse targets in images and tracking sparse frequencies in time-series estimation. In the context of frequency tracking, we illustrate the advantages of EMD-DF in tracking neural oscillations in electrophysiology recordings from rodent brains. We demonstrate that EMD-DF casually produces representations that achieve much higher time-frequency resolution than traditional causal linear methods such as the Short-Time Fourier Transform (STFT).

I. INTRODUCTION

Tracking algorithms (also called dynamic filtering) aim to improve the performance of statistical inference procedures for time series by incorporating information from a dynamics model which describes how the signal evolves. For example, the widely used Kalman filter [1] efficiently produces optimal estimates from linear measurements under additive Gaussian noise in the measurement and dynamics models. However, in contrast to these classic models, sparsity models are extremely non-Gaussian and have become increasingly popular due to their state-of-the-art performance in a variety of problems (for example in image processing [2] and compressive sensing [3]). Sparse inference problems with static data vectors have been studied in depth, resulting in many algorithmic advances and performance guarantees [4], [3], [5], [6], [7].

In the spirit of the Kalman filter, sparse tracking algorithms have also been introduced for dynamic filtering when the sparse signals are time-varying and have shown utility in practice [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20]. Basis pursuit denoising with dynamic filtering (BPDN-DF) [19] is one example of a recent algorithm which penalizes differences between the data and the prediction with an $\ell_p$-norm to incorporate a dynamics model into the regularization. However, in many applications with discretized domains, such $\ell_p$-norm regularizers disproportionately penalize predictions with slight mismatch in the signal support because they do not incorporate any knowledge of meaningful geometry (when it exists) into the penalty. Consider, for example, an imaging scenario where we wish to track a single pixel moving through a scene. An $\ell_p$-norm based regularizer assigns equal penalties to any prediction in which the target is not precisely in the correct location regardless of how far away the erroneous pixel is. Similarly, when tracking sparse frequency targets, an $\ell_p$-norm based regularizer on frequencies is agnostic to how similar the estimated frequency is to the frequency of interest.

In this work, we propose the earth mover’s distance (EMD) as an alternative regularizer for tracking time-varying sparse signals and introduce a new causal sparse tracking algorithm: earth mover’s distance dynamic filtering (EMD-DF). In essence, the EMD measures the amount of energy required to transform one signal into another, allowing the algorithm to account for relevant geometric relationships of the sparse coefficient space. The proposed EMD-DF is therefore most appropriate in situations where online estimation is necessary (e.g., closed-loop systems) and where there is a natural geometry to the coefficient space that should be respected (e.g., meaningful ordering). The main contribution of this paper is the introduction of EMD-DF at the algorithmic level, the casting of various versions (e.g., non-negative coefficients, complex-coefficients) of the problem at tractable numerical optimizations, and the formulation of highly-efficient approaches that reduce computational complexity for large-scale problems. Through a series of simulations, we demonstrate the advantages of EMD-DF compared to existing methods on tracking sparse targets in images and tracking sparse frequencies in time-series estimation. In the context of frequency tracking, we illustrate the advantages of EMD-DF in tracking neural oscillations in electrophysiology recordings from rodent brains. We demonstrate that EMD-DF casually produces representations that achieve much higher
We model the signal as evolving according to a dynamics governed by the matrix \( \parallel \cdot \parallel \) \( G \), one popular optimization-based approach is Basis-Pursuit Denoising Dynamic Filtering (BPDN-DF) which provides theoretical convergence guarantees, and Reweighted-\( \ell_1 \) Dynamic Filtering (RWL1-DF) which was found to be more robust to model mismatch [19]. We describe each of these algorithms in more detail as they provide intuition for the contributions of this paper. Furthermore, they offer state-of-the-art performance for causal algorithms, thus serving as points of comparison in the experiments to be described later.

BPDN-DF modifies standard BPDN with the addition of a tracking regularizer:

\[
\hat{x}_n = \text{argmin}_x \frac{1}{2} \| y - Ax \|^2 + \lambda \| x \|_1 + \gamma \| x - \tilde{x}_{n-1} \|_2, \tag{6}
\]

We may interpret this as the solution to a least-squares problem with the \( \ell_1 \)-norm as a sparsifying regularizer with parameter \( \lambda > 0 \) controlling the trade-off between measurement fidelity and signal sparsity. Reweighted-\( \ell_1 \) (RWL1) is one notable modification of BPDN that uses a hierarchical model called a Laplacian Scale Mixture to impose separate distributions on individual coefficients [30], [31]. Expectation maximization is then employed to estimate the signal, resulting in an iterative refinement of the signal via

\[
\hat{x}^k = \text{argmin}_x \| y - Ax \|^2 + \lambda_0 \sum_{i=1}^N \lambda_{k+1}\left[ i \right] \| x[i] \|_1, \tag{5}
\]

where \( k \) represents the algorithmic iteration, and \( \lambda_{k+1}\left[ i \right] = \frac{\beta}{\| x[i] \| + \eta} \). RWL1 can yield sparser solutions than BPDN, however this improved performance comes at the cost of solving multiple optimization programs.

While the sparse recovery techniques discussed above infer static sparse vectors, recent work has also extended these ideas to dynamic filtering for sparse time-varying signals. Early work in this area included batch (i.e., non-causal) approaches [32], [33], [34], [35], [36], [37], [38] and modifications to the causal Kalman filter [10], [11]. More recent causal approaches include Basis Pursuit Denoising Dynamic Filtering (BPDN-DF) which provides theoretical convergence guarantees, and Reweighted-\( \ell_1 \) Dynamic Filtering (RWL1-DF) which was found to be more robust to model mismatch [19]. We describe each of these algorithms in more detail as they provide intuition for the contributions of this paper. Furthermore, they offer state-of-the-art performance for causal algorithms, thus serving as points of comparison in the experiments to be described later.

A. Dynamic Filtering

Dynamic filtering is the problem of recovering a time varying signal from noisy measurements with the aid of a dynamics model. Here, we consider the linear observation model

\[
y_n = A_n x_n + \sigma \epsilon_n, \tag{1}
\]

where for each time step \( n \), \( x_n \) is the underlying signal, \( A_n \) is a linear observation operator, \( \sigma \epsilon_n \) is Gaussian measurement noise with variance \( \sigma^2 \), and \( y_n \) is the resulting measurement vector. We model the signal as evolving according to a dynamics function \( g \) as

\[
x_{n+1} = g(x_n) + \eta_n, \tag{2}
\]

where \( \eta_n \) is a noise vector called the innovations that accounts for inaccurate modeling of the dynamics. When \( g \) is linear and the signal, observation noise and innovations are Gaussian distributed, the classical Kalman filter provides an efficient way to compute the optimal (i.e., minimum expected \( \ell_2 \) error) estimate taking into account all measurements up to the current time step. The estimate produced by the Kalman filter may be expressed as

\[
\hat{x}_n = \text{argmin}_x \| y - A_n x \|^2_{R_n^{-1}} + \| x - G_n \hat{x}_{n-1} \|^2_{(Q_n + G_n P_{n-1} G_n^T)^{-1}}, \tag{3}
\]

where \( G_n \) is the linear dynamics operator, \( R_n \) and \( Q_n \) are covariance matrices of the noise and innovations, and \( \hat{x}_{n-1} \) and \( P_{n-1} \) are the previous signal estimate and its covariance. Here, we use \( \| \cdot \|_B \) to denote the norm induced by the positive-definite matrix \( B \) (i.e., \( \| a \|_B^2 = a^T B a \)). Thus, the Kalman filter may be interpreted as the solution to a least-squares problem which is regularized by the dynamics model. The Kalman filter and its extensions [23], [24] have been extensively used in many applications throughout science and engineering.

Sparsity models have also received much attention from the research community in recent years. A vector \( x \in \mathbb{C}^N \) is said to be sparse if only a few of its elements are non-zero (i.e., \( \| x \|_0 \ll N \), where \( \| \cdot \|_0 \) indicates the number of non-zero elements in the operand). Suppose \( y \) contains noisy observations of \( x \) through a linear measurement operator \( A \in \mathbb{C}^{M \times N} \). For example, results in the field of compressed sensing show that under certain conditions on \( A \), \( x \) may be recovered from \( y \) even when \( M \ll N \). Of the many sparse inverse algorithms that exist (e.g., [25], [5], [6], [26], [27], [28], [29]), one popular optimization-based approach is Basis-Pursuit Denoising (BPDN) [4]:

\[
\hat{x} = \text{argmin}_x \frac{1}{2} \| y - Ax \|^2 + \lambda \| x \|_1. \tag{4}
\]
Furthermore, this approach introduces additional algorithm parameters to tune and lowers the sparsity penalty in patches of elements around the active elements in the prediction, potentially resulting in a less sparse solution.

B. Earth Mover’s Distance

The earth mover’s distance (EMD) is a metric which grew out of the optimal transport (OT) literature initiated by Monge [40]. The EMD has recently been increasingly used in a variety of applications such as image and histogram comparison [41], [42], [43], as well as for sparse inverse problems [44], [45], [46]. We consider here the EMD for discrete signals. Intuitively, if we visualize the first signal as being composed of piles of dirt and the second as holes, the EMD computes the minimum amount of work needed to fill the holes with dirt. Formally, for two signals $x$ and $y$, the EMD solves

$$d_{emd}(x, y) = \min_F \sum_{ij} R_{ij} F_{ij}$$

subject to $F_{ij} \geq 0$,

$$\sum_j F_{ij} \leq x[i], \quad \sum_i F_{ij} \leq y[j],$$

$$\sum_{ij} F_{ij} = \min(||x||_1, ||y||_1),$$

(8)

where mass flow from the $i$-th element of $x$ to the $j$-th element of $y$, represented by the elements of the matrix $F = (F_{ij})$, incurs a cost given by $R_{ij}$. Often, this cost is defined as $R_{ij} = d(c_i, c_j)$ where $d$ is a distance metric and $c_i$ represents the discretized support coordinates of location $i$. The EMD is the cost associated with the minimum cost flow under four constraints. The first constraint specifies that flows must be positive. The second and third constraints enforce conservation of mass (e.g., the total mass flowing out of the $i$-th position of $x$ is bounded by $x[i]$). The final constraint states that the total amount of flow must be equal to the total mass of the smaller operand signal. This prevents the trivial solution where no mass flows, resulting in a cost of zero. A key property to note is that the EMD is inherently aware, by way of the distance matrix $(R_{ij})$, of the geometric relationship between elements in its operands. With non-trivial choices for this distance matrix, the cost to move mass to nearby elements is different than the cost to move mass over a longer distance. This is in stark contrast to $\ell_p$ metrics, and is the primary motivation for its use as a tracking regularizer.

The traditional EMD formulation presented here involves solving for $O(N^2)$ flow variables, which has the potential to be computationally prohibitive for large problems. While recent computational advances exploiting entropic regularization [47] enable fast numerical approaches to EMD problems (even in variational settings) [48], [49], these algorithms only approximate the EMD calculation and we defer consideration of these advances in tracking problems for future work. However, for applications whose distance cost $d$ is restricted to be the standard Euclidean distance (e.g., video), geometric structure can be exploited to also reduce the optimization variable complexity in exact EMD solutions from $O(N^2)$ to $O(N)$. In particular, the EMD problem can be reinterpreted as a fluid dynamics flux problem known as the Beckmann problem [50]. This problem searches for the optimal flux configuration of a fluid flowing between a source and a sink (i.e., the input arguments of the EMD problem). The enormous reduction in variables is therefore a result of physical fluid constraints that, by virtue of its representation, restricts point masses from “teleporting” across space. We note that this formulation has recently been applied [52] for computing the EMD between large-scale images.

In this work, we consider the discretization of a $D$-dimensional flux field $M \in \mathbb{R}^{N \times D}$, whose support is evenly gridded. For example, for images when $D = 2$, the columns of a 2-dimensional flux field (with equally gridded horizontal and vertical spacings) may be reorganized into two flux matrices $M_x, M_y \in \mathbb{R}^{n_x \times n_y}$, each one representing the flux field in each direction and where $N = n_x n_y$. The notion of how much each point in the flux field is a source or a sink is mathematically described by a linear divergence operator, defined as

$$\text{div}(M)[i, j] = (M_x[i, j] - M_x[i-1, j]) + (M_y[i, j] - M_y[i, j-1]),$$

where zero-flux boundary conditions are enforced (i.e., $M[i, j] = 0$ whenever $i$ or $j$ falls outside the support). Now, we can re-express the EMD definition in (8) as the Beckmann problem:

$$d_{emd}(x, y) = \min_M \|M\|_{2,1}$$

subject to $\text{div}(M) + y - x = 0$,  

(9)

where the rows of $M$ contain points in a $D$-dimensional vector field and $\|M\|_{2,1} := \sum_{i=1}^{N} \|m_i\|_2$ denotes the sum of their Euclidean norms. This optimization searches for the minimal vector field configuration $M$ whose inward and outward flux contributions are defined by $x$ and $y$ respectively.

While dramatically reducing the number of optimization variables for exact EMD calculation under Euclidean distance cost, a key limitation of the Beckmann formulation is that applications are limited to inputs that to lie in the probability simplex (i.e., vectors that sum to one). For example, in radar tracking, targets can spontaneously pop in and out, therefore the total energy is not constant over time. The trivial solution of normalizing their total energy (to fit these formulations) is a poor choice because individual signal energies will scale arbitrarily. To exploit the efficiency of (9), we will require a reformulation to adopt the type of constraints in (8) that more gracefully allows it to handle applications where the total energy changes with time.

III. EARTH MOVER’S DISTANCE DYNAMIC FILTERING

One drawback of existing tracking algorithms is a lack of robustness to small mismatches in the locations of the active signal coefficients, which is problematic when there is a geometric relationship or ordering among the coefficients. For instance, consider the image tracking scenario depicted in Figure 1. We should encourage signal estimates with active

---

1For more theoretical details and connections to optimal transport theory, see Villani’s excellent monograph [51] §1.2.3.
pixels geometrically close to the ground truth (even if the locations do not match exactly) and we should penalize estimates with active pixels that are far away. Unfortunately, each candidate estimate looks equally plausible when the error is measured with an $\ell_p$-norm on the difference vector (e.g., mean-squared error). Similarly, in the problem of tracking a set of time varying frequencies, the ordering of the frequencies in the DFT matrix results in a geometric relationship among the DFT coefficients which is not effectively utilized with $\ell_p$-norm regularizers. The EMD is a natural alternative regularizer in both of these scenarios.

We propose a new tracking algorithm, earth mover’s distance dynamic filtering (EMD-DF), where the causal estimate of the signal at time $n$ is given by:

$$\hat{x}_n = \arg\min_x \frac{1}{2} \|y_n - Ax\|_2^2 + \lambda \|x\|_1 + \gamma d_{\text{emd}}(x, \bar{x}_n),$$

where $\bar{x}_n = g(\bar{x}_{n-1})$ is the prediction from the previous time step. EMD-DF has a similar structural form as BPDN-DF at first glance, but the use of an EMD penalty instead of an $\ell_2$ dynamics regularizer is non-trivial because the evaluation of the EMD itself requires the solution of an optimization program. Incorporating the EMD into a dynamic filtering algorithm for common signals of interest presents three challenges that require technical innovation: the traditional formulation of the EMD 1) operates exclusively on non-negative vectors, 2) operates on real-valued vectors, and 3) requires a prohibitive computational complexity for inclusion inside an optimization program. We address each of these issues in the following subsections.

A. EMD-DF for Nonnegative Signals

For the case where the signal of interest is nonnegative, we can substitute the definition of the EMD into (10) and arrive at the following joint optimization over the signal estimate and the EMD flow variables $F$:

$$\tilde{x}_n = \arg\min_{x, F} \frac{1}{2} \|y_n - Ax\|_2^2 + \lambda \|x\|_1 + \gamma \sum_{ij} R_{ij} F_{ij}$$

subject to $F_{ij} \geq 0$,

$$\sum_j F_{ij} \leq |x[i]|, \sum_i F_{ij} \leq |\tilde{x}_n[j]|,$$

$$\sum_{ij} F_{ij} = \min (\|x\|_1, \|\tilde{x}_n\|_1).$$  

(11)

Here, we adopt the notation $\arg\min_{x, F} h(x, F) = \arg\min_x [\min_F h(x, F)]$. The last constraint is non-linear and thus complicates the evaluation of the optimization program. To address this challenge, we replace the nonlinear equality constraint by introducing a slack variable as follows:

$$\hat{x}_n = \arg\min_{x, F, u} \frac{1}{2} \|y_n - Ax\|_2^2 + \lambda \|x\|_1$$

$$+ \gamma \sum_{ij} R_{ij} F_{ij} - \mu u$$

subject to $F_{ij} \geq 0$,

$$\sum_j F_{ij} \leq |x[i]|, \sum_i F_{ij} \leq |\tilde{x}_n[j]|,$$

$$\sum_{ij} F_{ij} = u, \|x\|_1 \geq u, \|\tilde{x}_n\|_1 \geq u.$$  

(12)

The additional term in the objective function encourages the slack variable $u$ to be as large as possible, while the additional constraints force $u$ to be bounded above by $\|x\|_1$ and $\|\tilde{x}_n\|_1$. Hence, for an appropriate value of $\mu$, $u$ will be equal to $\min (\|x\|_1, \|\tilde{x}_n\|_1)$, as desired. Although $\mu$ is an additional parameter which must be tuned, we observed experimentally that performance is robust to the particular choice of $\mu$ over a substantial range.

B. EMD-DF for Complex-valued Signals

In some applications such as tracking in the frequency domain (e.g., DFT coefficients), the signal of interest is complex valued. We now expand the formulation of EMD-DF for nonnegative signals from the previous section to deal with complex valued signals. The natural modification would be to simply ignore the signal phase and constrain flows based on magnitudes. That is, we would like to solve

$$\tilde{z}_n = \arg\min_{z, F, u} \frac{1}{2} \|y_n - Az\|_2^2 + \lambda \|z\|_1$$

$$+ \gamma \sum_{ij} R_{ij} F_{ij} - \mu u$$

subject to $F_{ij} \geq 0$,

$$\sum_j F_{ij} \leq |z[i]|, \sum_i F_{ij} \leq |\tilde{z}_n[j]|,$$

$$\sum_{ij} F_{ij} = u, \|z\|_1 \geq u, \|\tilde{z}_n\|_1 \geq u.$$  

(13)
and easily solved via an off-the-shelf optimization package (e.g., CVX [53], [54] or TFOCS [55]). First, we decompose both the real and imaginary parts of \( z \) into positive and negative components such that

\[
\begin{align*}
  z = (z_\text{re}^+ - z_\text{re}^-) + i (z_\text{im}^+ - z_\text{im}^-),
\end{align*}
\]

where \( z_\text{re}^+, z_\text{re}^-, z_\text{im}^+, z_\text{im}^- \in \mathbb{R}^N_+ \). Ideally, we would like a decomposition in which the positive and negative components do not overlap, such as

\[
\begin{align*}
  z_\text{re}^+[i]z_\text{re}^-[i] = z_\text{im}^+[i]z_\text{im}^-[i] = 0, \quad i = 1, \ldots, N. \tag{14}
\end{align*}
\]

In this case, the magnitude of the real and imaginary parts may be evaluated simply by adding the corresponding positive and negative component vectors. We can then approximate the magnitude of each element as \( z_\text{re}^+[i] + z_\text{re}^-[i] + z_\text{im}^+[i] + z_\text{im}^-[i] \).

Since \( a^2 + b^2 \leq (a+b)^2 \leq 2(a^2 + b^2) \) for all \( a, b \in \mathbb{R}^+ \), this approximation is accurate within a factor of \( \sqrt{2} \). This ideal decomposition always exists; for example, it may be computed using the map \( a^+ = \max(a, 0) \) and \( a^- = -\min(a, 0) \).

The resulting convex relaxation of (13) is then given by

\[
\begin{align*}
  \hat{z}_n = \arg \min_{z^*, F_{ij}} & \frac{1}{2} \|y_n - A'z^*\|_2^2 + \lambda \|z^*\|_1 \\
  \text{subject to} & \quad F_{ij} \geq 0, \\
  & \quad \sum_i F_{ij} \leq \|z_n^-[j]\|_1 + \sum_j F_{ij} = u, \\
  & \quad \|z^*\|_1 \geq u, \quad \|\hat{z}_n\|_1 \geq u, \tag{15}
\end{align*}
\]

where,

\[
A' = [A -A \quad iA -iA],
\]

and \( z^* \) is the concatenation of the decomposed real and imaginary parts of \( z \). Note that the decomposition produced by this optimization is not guaranteed to satisfy (14). However, the \( \ell_1 \) regularizer serves to discourage solutions containing energy in overlapping elements of the positive and negative components.

C. EMD Computational Complexity

For general cost distances, the optimization program (12) involves solving \( N \) signal variables and an additional \( N^2 \) flow variables. Thus, the addition of the EMD regularizer potentially incurs a prohibitive increase in computational complexity compared to algorithms such as BPDN or RWL1. For general distance costs, note that when only \( K \) elements of \( x_\text{re} \) are non-zero the conditions \( F_{ij} \geq 0 \) and \( \sum_i F_{ij} \leq x_\text{re}[i] \) imply that all but \( K \) columns of \( F \) contain only zeros. Hence, regardless of the distance cost, we need only solve for \( NK \) flow variables for sparse signal tracking, resulting in significant savings in computational cost when \( K \ll N \).

Furthermore, in the common case when the distance cost \( d \) is Euclidean, we can exploit Beckmann’s formulation of the EMD (9) to reduce the number of EMD variables from \( O(N^2) \) to \( O(N) \). This formulation, however, requires that the signals have unit mass (i.e., \( \|x\|_1 = \|y\|_1 = 1 \)), meaning that we cannot simply apply the pre-existing method. In the following, we outline how a reformulation of the Beckmann problem for unequal total masses [56] may be incorporated into the EMD-DF program. To allow input arguments with unequal total mass, we introduce slack variables \( w, v \) to artificially bound the flux from the original source \( x \) and sink \( y \). The modified EMD program is then:

\[
\begin{align*}
  d_{\text{emd}}(x, y) = \min_{M, w, v} & \quad \|M\|_{2,1} \\
  \text{subject to} & \quad \text{div}(M) + v - w = 0, \\
  & \quad 0 \leq w \leq x, 0 \leq v \leq y, \\
  & \quad \|w\|_1 = \|v\|_1 = \min(\|x\|_1, \|y\|_1), \tag{16}
\end{align*}
\]

where \( w, v \) are nonnegative vectors with similar dimensions as \( x, y \). This optimization searches for the minimal vector field configuration that describes, via the first constraint, its flux to be travelling between a source \( w \) and a sink \( v \). The second constraint describes the source and sink as nonnegative slack variables that are bounded above by their proxies \( x \) and \( y \) respectively; this constraint is analogous to the mass preservation constraints in (8). The last constraint states that the induced flux must be bounded by the total mass of the smaller operand signal, which is similar in spirit to the fourth constraint of (8). This formulation has \( N(D+2) \) variables, where \( D \) is the dimensions of the vector field (e.g., \( D = 2 \) for images). Applying this EMD formulation, (12) becomes

\[
\begin{align*}
  \hat{x}_n = \arg \min_{x, M, u, v, \tilde{v}} & \quad \frac{1}{2} \|y_n - Ax\|_2^2 + \lambda \|x\|_1 \\
  + \gamma \|M\|_{2,1} - \mu u \\
  \text{subject to} & \quad \text{div}(M) + \tilde{v} - v = 0, \\
  & \quad 0 \leq v \leq x, \quad 0 \leq \tilde{v} \leq \tilde{x}, \\
  & \quad \|v\|_1 = \|\tilde{v}\|_1 = u, \\
  & \quad u \leq \|x\|_1, \quad u \leq \|\tilde{x}\|_1, \tag{17}
\end{align*}
\]

where we have introduced another slack variable \( u \) to linearize the minimum operator in (16) to make the program convex. The complex variant of EMD-DF given by (15) can also be trivially converted to adopt this formulation, though it is not shown here for the sake of brevity.

IV. RESULTS

In this section, we demonstrate the utility and performance of EMD-DF through a series of simulations on synthetic and real data. First, we consider an imaging scenario where the goal is to track targets moving throughout the scene. Next, we study the problem of tracking time-varying frequencies in a 1-D time series. We then use the same approach to track neural oscillations in electrophysiology data. Finally, we demonstrate the significant numerical speed up of EMD-DF due to Beckmann’s formulation for imaging applications.

Throughout these experiments, we use the Templates for First-Order Conic Solvers (TFOCS) [55] software package to solve the optimization problems for BPDN, BPDN-DF, RWL1,
and RWL1-DF, and the variant of EMD-DF for nonnegative signals. For the complex valued variant of EMD-DF, we use the CVX software package [53], [54].

A. Target Tracking

In this simulation, we consider an imaging experiment in which sparse targets move around a scene. Each image consists of 10 × 10 pixels where objects are represented with single active pixels and the remaining pixels are equal to zero. Rather than directly observing the images, we observe noisy linear measurements through a Gaussian observation operator (i.e., compressive sensing measurements). At each time step, targets move randomly to adjacent locations via discrete Brownian motion. An example trajectory is shown in Figure 2.

An example recovery for a single time point is shown in Figure 3. Because no information about the direction of object movement is available, the predictions for each algorithm are formed using an identity dynamics model (i.e., \( \mathbf{x}_n = \mathbf{x}_{n-1} \)). The EMD dynamics regularizer shows a clear qualitative benefit, even though the precise locations of the prediction do not align with those in the ground truth.

Next, we evaluate algorithm performance by quantitatively computing the relative mean-square-error (rMSE) defined by

\[
\text{rMSE} = \frac{\| \mathbf{x}_n - \mathbf{\hat{x}}_n \|_2^2}{\| \mathbf{x}_n \|_2^2}. \tag{18}
\]

Figure 4 shows the rMSE for the same simulated video snippets recovered using EMD-DF and a host of other sparse recovery algorithms. Note that EMD-DF maintains the lowest rMSE for the entire segment and the two competing tracking algorithms actually perform worse than BPDN (which does not account for the dynamics model at all). When \( \gamma = 0 \), BPDN-DF reduces to BPDN, however we use a modest positive value for \( \gamma \) to demonstrate how the dynamics mislead recovery when using the \( \ell_p \) norm as a regularizer. Next, the plots in Figure 5 demonstrate the mean performance of the various algorithms as functions of the sparsity level, \( K \), and as a function of the number of measurements taken, \( M \). Compared to the competing algorithms, EMD-DF is able to successfully track more targets for a given number of measurements, or successfully track a given number of targets using fewer measurements.

Finally, Figure 6 shows rMSE as a function of target movement speed. For low-speed targets, a static model with a small blurring kernel in the prediction may work well in conjunction with a traditional regularization approach using an \( \ell_p \) norm regularizer. However, this approach requires another parameter to be continually adapted to the current target speed and wider blurring kernels result in more information loss in the prediction. To illustrate these effects, we simulate target tracking over different speeds while incorporating a 3 × 3 averaging filter into the dynamics prediction for RWL1-DF. We see in this plot that target speed does affect the overall performance of all methods tested, but EMD-DF in general demonstrates more robustness to variations in target speed.

B. Frequency Tracking

In the next series of simulations, we study the performance of EMD-DF for complex valued signals via a frequency tracking task. In particular, we observe noisy measurements of a 1-D time series which is composed of \( K \) sinusoids of different frequencies that change as a function of time:

\[
y(t) = \sum_{k=1}^{K} a_k \cos(2\pi f_k(t) + \phi_k(t)) + \sigma \varepsilon(t). \tag{19}
\]

where the \( a_k \) are chosen by dividing the unit interval according to a uniform distribution, the phases \( \phi_k(t) \) are chosen deterministically based on the \( f_k(t) \) to ensure that \( y(t) \) is continuous, and the last term represents additive Gaussian noise with variance \( \sigma^2 \). The frequencies \( f_k(t) \) change every \( c_f \) samples by an amount \( c_f \), where \( c_t \sim N(\mu_t, \sigma_t) \) and \( c_f \sim N(0, \sigma_f) \). New values of \( c_t \) and \( c_f \) are drawn after each frequency change.
Fig. 4. Recovery performance over time. Top: the rMSE for estimates produced by various recovery algorithms (averaged over 20 trials) is plotted as a function of time. The original image size is $10 \times 10$ pixels, measurement vectors have length $M = 20$, and each frame contains $K = 5$ targets. EMD-DF is the top performer throughout. An identity dynamics function is used in all of the tracking algorithms, however only EMD-DF is able to effectively use the information from the predictions. In fact, the inappropriate tracking regularizer in BPDN-DF and RWL1-DF actually degrades performance. Bottom: median recovery rMSE. For each algorithm, the red line indicates the median rMSE with respect to time. Box boundaries indicate the 25% and 75% percentiles and whiskers indicate minimum and maximum values (excluding outliers indicated by red crosses). The rMSE of EMD-DF at each time point is tightly clustered around its median which is lower than competing algorithms.

Frequencies are constrained to reside in a specified band; if $c_f$ is generated such that $f_k(t) + c_f$ is outside of the specified band, $c_f$ is regenerated until a permissible value is produced. In the simulations that follow, data are generated with a sampling frequency of $f_s = 256$ Hz using $K = 3$, $\mu_t = 40$, $\sigma_t = 0$, $\sigma_f = 4$ and each frequency is banded between 0 and 128 Hz. Our goal is to recover denoised time-frequency plots with greater time and/or frequency resolution than is possible with standard short-time Fourier transform based methods. This can be accomplished within the sparse signal tracking framework by estimating sparse coefficients in an overcomplete DFT dictionary $\Phi$, where

$$\Phi_{mn} = \exp(i2\pi mn/N),$$

(20)

for $m = 0, \ldots, M - 1$ and $n = 0, \ldots, N - 1$. In this context, the values of $M$ and $N$ have a different interpretation than our previous experiments. The parameter $M$ controls the length of the analysis window. Larger values of $M$ provide lower noise and higher frequency resolution estimates at the expense of lower temporal resolution. The parameter $N$ controls the number of overcomplete DFT coefficients. Larger values of $N$ result in better frequency resolution in the dictionary, but a more challenging inference problem. We call the ratio $S = N/M$ the oversampling factor. The dynamics model used in these experiments is a simple denoising function $g(x) = \tau_q(x)$ which sets all but the largest $q$ elements of $x$ equal to zero. The value of $q$ is a parameter that controls the number of frequencies to track.

In the following experiments, we compare the performance of EMD-DF to other sparse recovery algorithms. We note that spectrogram reassignment is an alternative method for sharpening TF representations beyond what is possible with the standard STFT [57], [58], [59], [60], [61]. However, reassignment methods involve a batch procedure which reassigns energy in the spectrogram using a signal dependent transformation of the time-frequency plane. In contrast, EMD-DF is casual, a feature that is critical in online applications such as closed-loop control. Therefore, we do not provide a comparison to reassignment methods in these simulations.

In this task, we wish to evaluate the quality of spectral estimates directly in the frequency domain. However, the
Fig. 6. Algorithm performance as a function of target speed in target tracking imaging simulations for different levels of noise variance, \( \sigma^2 \). Images consist of \( 10 \times 10 \) pixels with \( K = 5 \) targets in the scene. Performance of BPDN-DF plummets when there is any support mismatch in the prediction due to its \( \ell_p \)-norm dynamics regularization term. By blurring the prediction, RWL1-DF is able to cope with small support mismatch between the prediction and the true signal. However, the addition of a blurring kernel introduces another parameter which may not be feasible to tune. Furthermore, blurring the prediction causes RWL1-DF to perform worse than BPDN-DF at high noise levels when there is no support mismatch. By contrast, EMD-DF handles more severe support mismatch with no additional parameters and its performance scales better as a function of noise level.

Ground truth, \( f_k(t) \), is continuous whereas the estimates take values on a gridded space defined by the analysis window, so we cannot compute rMSE directly as we did in previous experiments. To address this issue and to allow comparison of spectra with differing frequency resolution, each time slice of the spectrum estimate is upsampled to a common frequency grid with resolution exceeding any of the spectra to be compared. For each sample in time, the EMD is computed between the upsampled estimate and the ground truth. These distances are then summed to form an aggregate error. Figure 7 illustrates this computation for several example spectra.

Figure 8 shows recovery error as a function of measurement noise. Algorithm parameters (e.g. \( \lambda, \gamma, \) etc.) are tuned for each noise value via direct search [62]. We find that using parameters found via this method yield performance that matches or exceeds those found by manual tuning, a common practice in the evaluation of sparse recovery algorithms. The dynamics function used in EMD-DF is \( g(x) = \tau_q(x) \) (i.e., set all but the largest \( q \) elements equal to zero). The dynamics function used in RWL1-DF additionally blurs the estimate to approximate a local frequency preference in the inference. At low noise levels, the measurements are reliable enough that high accuracy recovery is possible without dynamics information, so all of the algorithms perform well. At exceedingly high noise levels, the predictions given by the dynamics model yield no additional information. In the middle region however, dynamics significantly aid recovery.

C. Tracking Neural Oscillations

In this section, we apply EMD-DF to the problem of spectrum estimation from neurophysiology recordings. Oscillatory behavior is prominent in neural recordings in a variety of settings and is thought to be a fundamental phenomenon in brain function. There is great interest in the neuroscience community to understand the functional role of these oscillations [63], [64].

In many studies, the tools used for spectral analysis of neural recordings are based on the classical short-time Fourier transform (STFT). The time and frequency resolution of such techniques is thus limited by the uncertainty principle which prevents simultaneously achieving high frequency and time resolution. Here, we study how higher TF resolution may be obtained by imposing a sparsity model on the data and using EMD-DF for recovery in an overcomplete DFT dictionary.

In particular, we study the phenomenon of oscillation phase coupling in the theta (4-7 Hz) and gamma (30-80 Hz) bands.
Fig. 7. Examples of error computation used in frequency tracking simulations. Spectral estimates \( \hat{s}_t \) are shown for a single time slice, and the ground truth at time \( t \) is indicated by \( s_t \). When the ground truth frequency (marked in green) is contained in the active bin in the spectral estimate (marked in blue), higher resolution estimates (upper-left) are favored over lower resolution estimates (upper-right). When the active bin does not contain the ground truth, estimates with center-of-mass closer to the ground truth are favored, regardless of the resolution of the estimate (bottom-left vs bottom-right).

Fig. 8. Mean spectral estimate error. Top: example of component frequencies and the resulting (noisy) time series data. Bottom: Mean spectral estimate error as a function of noise standard deviation \( \sigma \) for frequency tracking simulations. Observed signals consist of three frequencies which change randomly every 150 ms according to Brownian motion with standard deviation equal to 4Hz. The mean error across 1000 trials is shown. Error bars represent \( \alpha = 0.01 \) confidence intervals. For moderate noise levels, EMD-DF outperforms BPDN and RWL1-DF.

which is observed in tasks such as memory consolidation and learning of item-context associations [65], [66], [67], [68], [69]. We begin by generating synthetic data so that we have a ground truth against which to compare various sparse recovery algorithms. We generate data which consists of two components: a theta band frequency and a gamma band frequency which is modulated by that same theta frequency. More precisely, our simulated data are defined by

\[
y(t) = a_\theta \cos(2\pi f_\theta(t)t + \phi_\theta(t)) \\
\cdot [1 + a_\gamma \cos(2\pi f_\gamma(t)t + \phi_\gamma(t))] + \varepsilon(t),
\]  

where \( f_\theta(t) \) and \( f_\gamma(t) \) are theta and gamma band frequencies respectively that drift according to Brownian motion, \( a_\theta \) and \( a_\gamma \) are their respective amplitudes, and \( \varepsilon(t) \) is Gaussian noise.

The phases \( \phi_\theta(t) \) and \( \phi_\gamma(t) \) change at frequency change points to prevent discontinuities. In the simulations below, we choose \( a_\theta = 1 \) and \( a_\gamma = 0.2 \).

We use EMD-DF in the same way described in IV-B. In this setting, the power in the theta band is much greater than that in the gamma band, so recovery of the theta band component is trivial. Thus, we modify the error metric by masking out the theta band frequency to concentrate on the recovery of frequencies in the more challenging band.

First, we consider how performance scales as a function of window length, which directly determines time resolution. Recovery error is plotted as a function of window length in Figure 9. Compared to competing methods, EMD-DF produces estimates with lower error, especially when shorter window lengths are used. In this experiment, we keep the oversampling factor constant (\( S = 5 \)), so using a longer window length results in higher frequency resolution. Furthermore, including more observations in our analysis window results in lower noise estimates. Both of these factors outweigh the loss in temporal resolution, and error thus decreases as a function of window length.

Finally, we employ EMD-DF to estimate the spectrum in a segment of real electrophysiology data recorded from a tetrode in rat hippocampus [70]. We set the dynamics function to track the top two frequencies \( \varrho(x) = \tau_2(x) \), and use an oversampling factor of \( S = 5 \). Figure 10 shows TF plots produced by the spectrogram and EMD-DF. Because EMD-DF
Fig. 9. Mean frequency recovery error as a function of window length for simulated neural oscillation signals. Data is generated with a sampling frequency of 256 Hz and consists of a theta band (4-7 Hz) component and a gamma band (30 - 80 Hz) component which is modulated by the amplitude of the theta band activity. Frequencies of the theta and gamma band components change randomly according to Brownian motion every 150 ms with standard deviations of 0.5 Hz and 6 Hz respectively. Shown above is the mean error averaged over 300 trials. Error bars represent $\alpha = 0.01$ confidence intervals. All of the sparse recovery algorithms offer vastly improved performance compared to the STFT based spectrogram. EMD-DF is the top performer, especially for shorter window lengths where the inference problem is particularly difficult.

utilizes the overcomplete DFT matrix for recovery, it produces a TF plot with vastly improved frequency resolution. Additionally, the spectrogram suffers from severe leakage in the theta band frequencies, an artifact which is not present in the sparse TF representation. Finally, we note that the improved resolution of the sparse TF plot reveals more subtle oscillatory dynamics that cannot be observed in the spectrogram.

D. Computational Scalability

Given the increases in performance and robustness demonstrated by EMD-DF, we are especially interested in improving computational complexity so that the algorithm can still scale well in practical applications with large state spaces. Here we examine the impact of adapting an approach based on Beckmann’s EMD formulation into our tracking problem. We note that frequency tracking (from the previous sections) also benefits from this formulation because it can be treated as a image tracking problem (spectrograms being images of single-pixel widths).

We conducted a similar simulation detailed in section IV-A and scaled the problem between image sizes of $12 \times 12$ ($N = 144$) and $48 \times 48$ ($N = 2304$), where $N$ is the number of pixels. For each image size, the pixel sparsity level was fixed at 5%. Each experiment was repeated 10 times for statistics aggregation and error bars denote $\pm 1$ standard deviation from the mean. Because our major concern is whether the proposed computational modification degrades EMD-DF performance over the general (but expensive) formulation, we compare differences between solutions using the root mean squared error (RMSE): $\sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - y_i)^2}$, where $x_i$ and $y_i$ are individual pixel intensity values of the respective solutions. We use the CVX software package (employing interior point methods) for both formulations for a fair comparison and measure relative runtime on a personal computer (Intel Core i7 with 3.5 GHz processor speed).

Figure 11 shows that by using the Beckmann’s formulation of EMD-DF (17), we obtain a significant speed up over the general formulation of EMD-DF (12). Figure 11 also shows that the difference between the Beckmann’s formulation of EMD-DF and the general formulation have very small differences. Taken together, these results demonstrate that the proposed re-formulation is much more computationally tractable and scalable to larger problem sizes while producing solutions that are essentially the same as the general approach.

V. SUMMARY AND FUTURE WORK

The estimation of signals that traverse a gridded domain can be enhanced by regularizing for underlying sparsity and dynamical structure. While current tracking methods in the literature have investigated sparsity in a number of ways, the issue of dynamical support mismatch remains a challenging open problem. To address this, we apply the EMD as a tracking regularizer for time varying sparse inverse problems in our proposed EMD-DF algorithm. The EMD provides a natural
We compare the runtime and the difference in solutions for two formulations (e.g., for perturbation experiments or closed-loop control). A method that reduces the problem to require only $O(N)$ that a reformulation of EMD yields an extremely efficient burden can be prohibitive. When transport distances are tracking using the EMD as a regularizer, this computational involve $\text{DF}$ is a causal algorithm making it applicable for online systems time-frequency plots such as spectrogram reassignment, EMD- In contrast to other approaches for increasing the readability of representations with resolution in both time and frequency that show how EMD-DF can be used to produce time-frequency tracking. We empirically that both variants outperform competing sparse complex valued signals (as a convex relaxation) and show this requires a reformulation such that it fits into a natural geometric framework that specifically computes the amount of support mismatch between two signals measured over a fixed-grid. However, since the EMD is itself an optimization problem, this requires a reformulation such that it fits into a natural setup for sparse inverse problems. In this work, we introduce two convex algorithms for tracking nonnegative signals and complex valued signals (as a convex relaxation) and show empirically that both variants outperform competing sparse recovery algorithms. In the context of frequency tracking, we show how EMD-DF can be used to produce time-frequency representations with resolution in both time and frequency that exceed what is possible with traditional methods like the STFT. In contrast to other approaches for increasing the readability of time-frequency plots such as spectrogram reassignment, EMD-DF is a causal algorithm making it applicable for online systems (e.g., for perturbation experiments or closed-loop control).

Computations using the traditional formulation of the EMD involve $O(N^2)$ flow variables. In the context of real time tracking using the EMD as a regularizer, this computational burden can be prohibitive. When transport distances are Euclidean (such as in image-tracking applications), we show that a reformulation of EMD yields an extremely efficient method that reduces the problem to require only $O(N)$ optimization variables. This recasting of the problem of interest into a more efficient optimization program dramatically reduces computational complexity to allow EMD-DF to be run efficiently for non-trivial problem sizes.

The EMD calculation may remain prohibitive for extremely large problems or for more general cases that do not use Euclidean distances. Fortunately, recent work in the optimal transport literature studies methods for more efficient computation of the EMD using a variety of relaxation techniques. Future work will focus on algorithmic advances [47], [48] to incorporate these techniques into the problem of sparse signal tracking.

ACKNOWLEDGEMENT
The authors would like to thank C. Kemere for the tetrode recording data.

REFERENCES

[1] R. E. Kalman, “A new approach to linear filtering and prediction problems,” Journal of Basic Engineering, vol. 82, no. 1, pp. 35–45, Mar. 1960.
[2] M. Elad, M. A. T. Figueiredo, and Y. Ma, “On the role of sparse and redundant representations in image processing,” Proceedings of the IEEE, vol. 98, no. 6, pp. 972–982, Jun. 2010.
[3] R. G. Baraniuk, “Compressive sensing [lecture notes],” IEEE Signal Processing Magazine, vol. 24, no. 4, pp. 118–121, Jul. 2007.
[4] D. L. Donoho and J. Tanner, “Sparse nonnegative solution of underdetermined linear equations by linear programming,” Proceedings of the National Academy of Sciences of the United States of America, vol. 102, no. 27, pp. 9446–9451, May 2005.
[5] J. A. Tropp and A. C. Gilbert, “Signal recovery from random measurements via orthogonal matching pursuit,” IEEE Transactions on Information Theory, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.
[6] D. Needell and J. A. Tropp, “CoSaMP: Iterative signal recovery from incomplete and inaccurate samples,” Applied and Computational Harmonic Analysis, vol. 26, no. 3, pp. 301–321, May 2009.
[7] S. Foucart and H. Rauchut, A Mathematical Introduction to Compressive Sensing, ser. Applied and numerical harmonic analysis. New York: Birkhäuser, 2013.
[8] M. S. Asif and J. Romberg, “Dynamic updating for $\ell_1$ minimization,” IEEE Journal of Selected Topics in Signal Processing, vol. 4, no. 2, pp. 421–434, Apr. 2010.
[9] M. S. Asif, A. Charles, J. Romberg, and C. Rozell, “Estimation and dynamic updating of time-varying signals with sparse variations,” in 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), May 2011, pp. 3908–3911.
[10] A. Carmi, P. Gurfil, and D. Kanevsky, “Methods for sparse signal recovery using Kalman filtering with embedded pseudo-measurement norms and quasi-norms,” IEEE Transactions on Signal Processing, vol. 58, no. 4, pp. 2405–2409, Apr. 2010.
[11] N. Vaswani, “Kalman filtered compressed sensing,” in 2008 15th IEEE International Conference on Image Processing, Oct. 2008, pp. 893–896.
[12] D. Zachariah, S. Chatterjee, and M. Jansson, “Dynamic iterative pursuit,” IEEE Transactions on Signal Processing, vol. 60, no. 9, pp. 4967–4972, Sep. 2012.
[13] J. Ziniel, L. C. Potter, and P. Schniter, “Tracking and smoothing of time-varying sparse signals via approximate belief propagation,” in 2010 Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers, Nov. 2010, pp. 808–812.
[14] N. Vaswani, “LS-CS-residual (LS-CS): Compressive sensing on least squares residual,” IEEE Transactions on Signal Processing, vol. 58, no. 8, pp. 4108–4120, Aug. 2010.
[15] 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, Sep. 2010.
[16] N. Vaswani, “LS-CS-residual (LS-CS): Compressive sensing on least squares dynamical systems,” in 2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton), Sep. 2010, pp. 1730–1736.
[62] T. Kolda, R. Lewis, and V. Torczon, “Optimization by direct search: New perspectives on some classical and modern methods,” *SIAM Review*, vol. 45, no. 3, pp. 385–482, Jan. 2003.

[63] L. Cornelissen, S.-E. Kim, P. L. Purdon, E. N. Brown, and C. B. Berde, “Age-dependent electroencephalogram (EEG) patterns during sevoflurane general anesthesia in infants,” *eLife*, vol. 4. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4502759/

[64] J. Aru, J. Aru, V. Priesemann, M. Wibral, L. Lana, G. Pipa, W. Singer, and R. Vicente, “Untangling cross-frequency coupling in neuroscience,” *Current Opinion in Neurobiology*, vol. 31, pp. 51–61, Apr. 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0959438814001640

[65] B. Schack, N. Vath, H. Petsche, H. G. Geissler, and E. Möller, “Phase-coupling of theta–gamma EEG rhythms during short-term memory processing,” *International Journal of Psychophysiology*, vol. 44, no. 2, pp. 143–163, May 2002.

[66] R. T. Canolty, E. Edwards, S. S. Dalal, M. Soltani, S. S. Nagarajan, H. E. Kirsch, M. S. Berger, N. M. Barbaro, and R. T. Knight, “High gamma power is phase-locked to theta oscillations in human neocortex,” *Science*, vol. 313, no. 5793, pp. 1626–1628, Sep. 2006.

[67] O. Jensen and L. L. Colgin, “Cross-frequency coupling between neuronal oscillations,” *Trends in Cognitive Sciences*, vol. 11, no. 7, pp. 267–269, Jul. 2007.

[68] A. B. L. Tort, R. W. Komorowski, J. R. Manns, N. J. Kopell, and H. Eichenbaum, “Theta–gamma coupling increases during the learning of item–context associations,” *Proceedings of the National Academy of Sciences*, vol. 106, no. 49, pp. 20942–20947, Aug. 2009.

[69] M. A. Belluscio, K. Mizuseki, R. Schmidt, R. Kempter, and G. Buzsáki, “Cross-frequency phase–phase coupling between theta and gamma oscillations in the hippocampus,” *Journal of Neuroscience*, vol. 32, no. 2, pp. 423–435, Jan. 2012.

[70] C. Kemere, M. F. Carr, M. P. Karlsson, and L. M. Frank, “Rapid and continuous modulation of hippocampal network state during exploration of new places,” *PLOS ONE*, vol. 8, no. 9, p. e73114, Sep. 2013.