Abstract—We consider the problem of spatiotemporal sampling in an evolutionary process \( x^{(n)} = A^n x \) where an unknown operator \( A \) driving an unknown initial state \( x \) is to be recovered from a combined set of coarse spatial samples \( \{x|_{\Omega_0}, x^{(1)}|_{\Omega_1}, \cdots, x^{(N)}|_{\Omega_N}\} \). In this paper, we will study the case of infinite dimensional spatially invariant evolutionary process, where the unknown initial signals \( x \) are modeled as \( \ell^2(\mathbb{Z}) \) and \( A \) is an unknown spatial convolution operator given by a filter \( a \in \ell^2(\mathbb{Z}) \), so that \( Ax = a \ast x \). We show that \( \{x|_{\Omega_m}, x^{(1)}|_{\Omega_m}, \cdots, x^{(N)}|_{\Omega_m} : \Omega_m : N \geq m-1, \Omega_m = m\mathbb{Z}\} \) contains enough information to recover the Fourier spectrum of a typical low pass filter \( a \), if \( x \) is from a dense subset of \( \ell^2(\mathbb{Z}) \). The idea is based on a nonlinear, generalized Prony method similar to [2].

We are going to introduce the notion of infinite dimensional spatially invariant evolutionary system and uniform spatiotemporal sampling problem. Let \( x \in \ell^2(\mathbb{Z}) \) be an unknown initial spatial signal and the evolution operator \( A \) be given by an unknown convolution filter \( a \in \ell^2(\mathbb{Z}) \) such that \( Ax = a \ast x \). At time \( t = n \in \mathbb{N} \), the signal \( x \) has evolved to \( x_n = A^n x = a^n \ast x \), where \( a^n = a \ast a \cdots \ast a \). We call this evolutionary system spatially invariant. In this paper, we are interested in the recovery of unknown filter \( a \) that drives the evolutionary process. Without loss of generality, let \( m \) be a positive odd integer \( (m > 1) \) and denote by \( S_m : \ell^2(\mathbb{Z}) \rightarrow \ell^2(\mathbb{Z}) \) the sampling operator on \( \Omega_m = m\mathbb{Z} \), i.e., \( (S_m x)(k) = x(mk) \). At time level \( t = l \), we have partial observations

\[
y_l = S_m(a^l \ast x)
\]

The special case we are going to consider can be stated as follows:

Under what conditions on \( a, m, N, x \) can \( a \) be recovered from the spatiotemporal samples \( \{y_l : l = 0, \cdots, N-1\} \), or equivalently, from \( \{x|_{\Omega_m}, (a \ast x)|_{\Omega_m}, \cdots, (a^{N-1} \ast x)|_{\Omega_m}\} \)?

In [2], Aldroubi and Kristal consider the recovery of unknown \( d \times d \) matrix \( B \) and unknown initial state \( x \in \ell^2(\mathbb{Z}_d) \) from coarse spatial samples of its successive states \( \{B^k x, k = 0, 1, \cdots\} \). Given an initial sampling set \( \Omega \subset \mathbb{Z}_d = \{1, \cdots, d\} \), they employ a new method related to Krylov subspace methods to show how large \( l_i \) should be to recover all the eigenvalues of \( B \) that can possibly be recovered from spatiotemporal samples \( \{B^k x(i) : i \in \Omega, k = 0, 1, \cdots, l_i - 1\} \). Our setup is very

I. INTRODUCTION

A. The dynamical sampling problem

In situations of practical interest, physical systems evolve in time under the action of well-behaved operators such as diffusion processes. Sampling of such an evolving system is done by sensors or measurement devices that are placed at various locations and can be activated at different times. For practical reasons, we aim to reconstruct any states in the evolutionary process using as few sensors as possible, but allow one to take samples at different time levels. This setting has not been studied within the classical approach in sampling theory, where the samples are taken simultaneously at only one time level. Dynamical sampling is a newly proposed sampling framework. It involves studying the time-space patterns formed by the locations of the measurement devices and the times of their activation. Mathematically speaking, suppose \( x \) is an initial distribution that is evolving in time satisfying the evolution rule:

\[
x_t = A_t x
\]

where \( \{A_t\}_{t \in [0, \infty)} \) is a family of evolution operators satisfying the condition \( A_0 = I \). Dynamical sampling asks the question: when do coarse samplings taken at varying times \( \{x|_{\Omega_0}, (A_t x)|_{\Omega_1}, \cdots, (A_N x)|_{\Omega_N}\} \) contain the same information as a finer sampling taken at the earliest time? One goal of dynamical sampling is to find all spatiotemporal sampling sets \( (\chi, \tau) = (\Omega_t, t \in \tau) \) such that certain classes of signals \( x \) can be recovered from the spatiotemporal samples \( x_t(\Omega_t), t \in \tau \).

In the above cases, the evolution operators are assumed to be known. It has been well-studied in the context of various evolutionary systems in a very general setting, see [3], [12], [8], [9], [10].

Another important problem arises when the evolution operators are themselves unknown or partially known. In this case, we are interested in finding all spatiotemporal sampling sets and certain classes of evolution operators so that the family \( \{A_t\}_{t \in [0, \infty)} \) or their spectrum can be identified. We call such a problem the unsupervised system identification problem in dynamical sampling.

B. Problem Statement

We are going to introduce the notion of infinite dimensional spatially invariant evolutionary system and uniform spatiotemporal sampling problem. Let \( x \in \ell^2(\mathbb{Z}) \) be an unknown initial spatial signal and the evolution operator \( A \) be given by an unknown convolution filter \( a \in \ell^2(\mathbb{Z}) \) such that \( Ax = a \ast x \). At time \( t = n \in \mathbb{N} \), the signal \( x \) has evolved to \( x_n = A^n x = a^n \ast x \), where \( a^n = a \ast a \cdots \ast a \). We call this evolutionary system spatially invariant. In this paper, we are interested in the recovery of unknown filter \( a \) that drives the evolutionary process. Without loss of generality, let \( m \) be a positive odd integer \( (m > 1) \) and denote by \( S_m : \ell^2(\mathbb{Z}) \rightarrow \ell^2(\mathbb{Z}) \) the sampling operator on \( \Omega_m = m\mathbb{Z}, \) i.e., \( (S_m x)(k) = x(mk) \). At time level \( t = l \), we have partial observations

\[
y_l = S_m(a^l \ast x)
\]
similar to the special case of regular invariant dynamical sampling problem in [2]. In this special case, they employ a generalization of the well-known Prony method that uses these regular undersampled spatiotemporal data first for the recovery of the Fourier spectrum of the correlating filter. Since the filter is a typical low pass filter with symmetry and monotonicity condition, it is completely determined by its Fourier spectrum. By using techniques developed in [10], one can recover the initial state. In this paper, we will address the infinite dimensional analog of this special case.

The remainder of the paper is organized as follows: In section II, we discuss the noise free case, and propose a generalized Prony method to show that we can reconstruct a via spatiotemporal samples \( \{y_i\}_{i=1}^{N} \), provided \( N \geq 2m - 1 \). In section III, we provide accuracy analysis of the generalized Prony method and the estimation results are formulated in the rigid \( \ell^\infty \) norm. In section IV, we do several numerical simulations to verify some estimation results. Finally, we summarize the work in section V.

C. Notations

Let us introduce some relevant notations. Let \( M = (m_{ij}) \) be a \( n \times n \) matrix, the infinity norm of \( M \) is defined by

\[
\|M\|_\infty = \max_{1 \leq i \leq n} \left( \sum_{j=1}^{n} |m_{ij}| \right)
\]

For a vector \( z = (z_i) \in \mathbb{C}^n \), we define the infinity norm \( \|z\|_\infty = \max_{1 \leq i \leq n} |z_i| \). It is easy to see that

\[
\|M\|_\infty = \max_{z \in \mathbb{C}^n, \|z\|_\infty=1} \|Mz\|_\infty
\]

We use \( z^T \) and \( M^T \) to denote their nonconjugate transpose.

II. Noise-free recovery

We consider the recovery of a frequently encountered case in applications when the filter \( a \in \ell^1(\mathbb{Z}) \) is a typical low pass filter and \( \hat{a}(\xi) \) is real, symmetric, and strictly decreasing on \([0, \frac{1}{2}]\). The symmetry reflects the fact that there is often no preferential direction for physical kernels and monotonicity is a reflection of energy dissipation. Without loss of generality, we also assume \( a \) is a normalized filter, i.e., \( |\hat{a}(\xi)| \leq 1, \hat{a}(0) = 1 \). Let \( \mu \) denote the Lebesgue measure on \( \mathbb{T} \), and \( X \) be a subclass of \( \ell^2(\mathbb{Z}) \) defined by

\[
X = \{ x \in \ell^2(\mathbb{Z}) : \mu(\{ \xi \in \mathbb{T} : \hat{x}(\xi) = 0 \}) = 0 \}
\]

Clearly, \( X \) is a dense class of \( \ell^2(\mathbb{Z}) \) under the norm topology. In noise free scenario, our first result shows, provided our initial state \( x \in X \), that we can completely determine \( a \) via a series of consecutive uniform undersampled states.

**Theorem 1.** Let \( x \in X \) be the initial signal and the evolution operator \( A \) be a convolution operator given by \( a \in \ell^1(\mathbb{Z}) \) so that \( \hat{a}(\xi) \) is real, symmetric, and strictly decreasing on \([0, \frac{1}{2}]\). Then \( a \) can be exactly recovered from measurements \( y_0, \ldots, y_{2m-1} \) defined in (1).

**Proof.** We show that the regular subsampled data \( \{y_i\}_{i=0}^{2m-1} \) contains enough information to recover the Fourier spectrum of \( a \) on \( \mathbb{T} \) up to a measure zero set. It is easy to show there exists a measurable set \( E \) with \( \mu(E) = 1 \) such that for each \( \xi \in E \), \( \{\hat{x}(\frac{\xi}{m}), \ldots, \hat{x}(\frac{\xi-2m+1}{m})\} \) is non-vanishing. For each \( \xi \in E \setminus \{0, \frac{1}{2}\} \), define the Hankel matrix

\[
H_m(\xi) = \left( \begin{array}{cccc}
\hat{y}_0(\xi) & \hat{y}_1(\xi) & \cdots & \hat{y}_{m-1}(\xi) \\
\hat{y}_1(\xi) & \hat{y}_2(\xi) & \cdots & \hat{y}_{m}(\xi) \\
\vdots & \vdots & \ddots & \vdots \\
\hat{y}_{m-1}(\xi) & \hat{y}_{m}(\xi) & \cdots & \hat{y}_{2m-2}(\xi) 
\end{array} \right)
\]

and the vector \( b_m(\xi) = (\hat{y}_m(\xi), \ldots, \hat{y}_{2m-1}(\xi))^T \in \mathbb{C}^m \), we can show that there is a unique \( q(\xi) = (q_0(\xi), \ldots, q_{m-1}(\xi))^T \in \mathbb{C}^m \) such that

\[
H_m(\xi)q(\xi) = b_m(\xi)
\]

Theorem 1 addresses the infinite dimensional analog of Theorem 4.1 in [2]. Once \( a \) is recovered, we can recover \( x \) using techniques developed in [8]. In general, for any \( a \in \ell^1(\mathbb{Z}) \), one can show the recovery of range of \( \hat{a} \) on the Torus except for a measure zero set only with minor modifications of the above proof.

**Definition 1.** Let \( a = (a(n))_{n \in \mathbb{Z}} \), the support set of \( a \) is defined by \( \text{Supp}(a) = \{ k \in \mathbb{Z} : a(k) \neq 0 \} \). If \( \text{Supp}(a) \) is a finite set, \( a \) is said to be a finite impulse response filter.

The term finite impulse response arises because the filter output is computed as a weighted, finite term sum, of past, present, and perhaps future values of the filter input. If \( a \) is a finite impulse response filter with support contained in \( \{-r, -r+1, \ldots, r\} \), we can get the following results immediately.

**Corollary 1.** In addition to the assumptions of Theorem 1, if \( a \) is a finite impulse response filter supported in \( \{-r, -r+1, \ldots, r\} \), then it is enough to recover \( \{\hat{a}(n) : i = 1, \ldots, r\} \) at \( r \) distinct locations via equation (3).

The proof of Theorem 1 doesn’t give a practical method in general, since it involves computing the Fourier transformation of infinite sequences and solving the roots of uncountably many polynomials. However, if we know in prior \( \text{Supp}(a) \) and \( \text{Supp}(x) \) are contained in \( \{-r, -r+1, \ldots, r\} \) for some \( r \in \mathbb{N}^+ \), then \( \{y_i\}_{i=0}^{2m-1} \) consists of sequences supported in \( \{-2mr, \ldots, 2mr\} \). We are able to compute \( \{\hat{y}_i(\xi)\}_{i=0}^{2m-1} \) for any \( \xi \in \mathbb{T} \). In this case, the proof of Theorem 1 essentially provides an algorithm for the recovery of the Fourier spectrum of \( a \). We summarize it as follows:
Algorithm II.1. Input: Choose $\xi \in \mathbb{T} - \{0, \frac{1}{2}\}$

1) From measurements $\{y_l\}_{l=0}^{2m-1}$, compute $\{\hat{y}_l(\xi)\}_{l=0}^{2m-1}$. Form the Hankel matrix $H_m(\xi)$ and $b_m(\xi)$ as in (2). Test the rank of $H_m(\xi)$, if $\text{rank}(H_m(\xi)) = m$, solve $q(\xi)$ via the following equation.

$$H_m(\xi) q(\xi) = -b_m(\xi)$$

2) For $q(\xi)$, form the Prony polynomial $p^e[z] = z^m + \sum_{l=0}^{m-1} q_l(\xi) z^l$. Find roots of $p^e[z]$ and order them to get $\{\hat{a}(\xi m^i) : i = 0, \ldots, m - 1\}$.

Output: $\{\hat{a}(\xi m^i) : i = 0, \ldots, m - 1\}$.

Remark 1. Note in this case, $H_m(\xi)$ is not invertible only at finitely many locations of $\mathbb{T}$, $H_m(\xi)$ is invertible with probability $1$.

Let $\{\xi_i : i = 1, \ldots, K\}$ be a subset of $\mathbb{T}$ satisfying the condition $|\xi_i - \xi_j| \neq \frac{k}{m}$ for $k = 0, \ldots, m - 1$, and $K m > r$. Assume we have recovered $\{\hat{a}(\xi m^i) : i = 0, \ldots, m - 1, j = 1, \ldots, K\}$ via Algorithm II.1, by Corollary 1, we can completely determine $a$.

III. Accuracy Analysis

In previous sections, we show that if we are able to compute the spectral data $\{\hat{y}_l(\xi)\}_{l=0}^{2m-1}$ at $\xi$, we can recover the Fourier spectrum $\{a(\xi m^i) : i = 0, \ldots, m - 1\}$ by Algorithm II.1. However, a critical issue remains. We need to analyze the accuracy of solution achieved by Algorithm II.1. The motivation to study the accuracy comes from two aspects. On one hand, for the case when one or both of $a$ and $x$ are not compactly supported, although we only have access to a finite section of each exact measurement $y_l \in \ell^2(\mathbb{Z})$ for practical reasons, we may have a good approximation $\hat{y}_l$ of $y_l$, so that $||\hat{y}_l(\xi) - y_l(\xi)||_{\infty} \leq \epsilon_i \ll 1$. Consequently, we can employ Algorithm II.1 to compute an approximation of the Fourier spectrum of $a$. A natural question to ask is how large the error will be between the approximate solutions and the actual solutions. On the other hand, even for the case when both $x$ and $a$ are compactly supported, what if we have noise in the process of computing $\{\hat{y}_l(\xi)\}_{l=0}^{2m-1}$? We can summarize our accuracy analysis problem in the following:

Assume the measurements are given by $\{\hat{y}_l(\xi)\}_{l=0}^{2m-1}$ compared to (1) so that $||\hat{y}_l(\xi) - y_l(\xi)||_{\infty} \leq \epsilon_i$ for all $\xi \in \mathbb{T}$. Given an estimate $\epsilon = \max_{l} |\epsilon_i|$, how large can the error in the reconstructed parameters in Step I and Step II of Algorithm II.1 be in the worst case in terms of $\epsilon$, and the true parameters.

Our accuracy analysis will consist of two steps. Suppose our measurements are perturbed from $\{\hat{y}_l(\xi)\}_{l=0}^{2m-1}$ to $\{\tilde{y}_l(\xi)\}_{l=0}^{2m-1}$. For any $\xi$, we first measure the perturbation of $q(\xi)$ in terms of $\ell^\infty$ norm. This step is linear and standard. Then we measure the perturbation of the roots. It is well known that the roots of a polynomial are continuously dependent on the small change of its lower degree coefficients. Hence, for a small perturbation, although the roots of the perturbed polynomial $p^e$ may not be real, we can order them according to their modulus and have a one to one correspondence with the roots of $p^e$.

Definition 2. Let $\xi \in \mathbb{T} - \{0, \frac{1}{2}\}$, consider the set $\{\hat{a}(\xi m^i) : i = 0, \ldots, m - 1\}$ that consists of $m$ distinct nodes.

1) For $0 \leq k \leq m - 1$, the “distance” between $\hat{a}(\xi m^i)$ with other $m - 1$ nodes is measured by the quantity

$$\delta_k(\xi) = \frac{1}{\prod_{0 \leq j \leq m-1} |\hat{a}(\xi m^j) - \hat{a}(\xi m^k)|}$$

2) For $0 \leq k \leq m$, the $k$th elementary symmetric function generated by the $m$ nodes is denoted by $\sigma_k(\xi) = \sum_{0 \leq j < \ldots < j_k \leq m-1} \hat{a}(\xi m^j)$, if $k = 0$, otherwise.

Proposition 1. Let the perturbed measurements $\{\tilde{y}_l(\xi)\}_{l=0}^{2m-1}$ be given with error satisfying $||\hat{y}_l(\xi) - \tilde{y}_l(\xi)||_{\infty} \leq \epsilon_i$ for all $l$. Let $\tilde{H}_m(\xi)$ and $\tilde{b}_m(\xi)$ be formed by $\{\tilde{y}_l(\xi)\}_{l=0}^{2m-1}$ in the same way as in (2). Assume $H_m(\xi)$ is invertible and $\epsilon$ is sufficient small so that $\tilde{H}_m(\xi)$ is also invertible. Denote by $\bar{q}(\xi)$ the solution of $\tilde{H}(\xi)\bar{q}(\xi) = \tilde{b}(\xi)$. Form the Prony polynomial $\tilde{p}^e$ by $\bar{q}(\xi)$ and let $\{\hat{a}(\xi m^i) : i = 0, \ldots, m - 1\}$ be its roots, then we have the following estimates as $\epsilon \to 0$.

$$||\bar{q}(\xi) - \hat{q}(\xi)||_{\infty} < ||H_m^{-1}(\xi)||_{\infty}(1 + m\beta_1(\xi))\epsilon + O(\epsilon^2)$$

where $\beta_1(\xi) = \max_{k=1, \ldots, m} |\sigma_k(\xi)|$. As a result, we achieve the following first order estimation

$$||\hat{a}(\xi m^i + \frac{i}{m}) - \bar{a}(\xi m^i + \frac{i}{m})|| < C_1(\xi)(1 + m\beta_1(\xi))||H_m^{-1}(\xi)||_{\infty}\epsilon + O(\epsilon^2)$$

where $C_1(\xi) = \delta_i(\xi) \cdot (\sum_{k=0}^{m-1} |\hat{a}(\xi m^i)|)$.

Therefore it is important to understand the relation between the behavior of $||H_m^{-1}(\xi)||_{\infty}$ and our system parameters, i.e., $m$, $\alpha$, and $x$. Next, we are going to estimate $||H_m^{-1}(\xi)||_{\infty}$ and reveal their connection with the spectral properties of $a$ and $x$.

Theorem 2. Assume $H_m(\xi)$ is invertible, then we have the lower bound estimation

$$||H_m^{-1}(\xi)||_{\infty} \geq m^2 \max_{i=0, \ldots, m-1} \frac{\beta_2(i, \xi)\delta_i(\xi)}{|x(\xi m^i + \frac{i}{m})|}$$

where $\beta_2(i, \xi) = \max_{k=0, \ldots, m-1} |\sigma_k(\xi)|$, and the upper bound estimation
Corollary 2. If \( |\tilde{x}(\xi)| \leq M \) for every \( \xi \in \mathbb{T} \), then
\[
\|H_{m^{-1}}^{-1}(\xi)\|_{\infty} \to \infty \text{ as } m \to \infty. \text{ In fact, } \|H_{m^{-1}}^{-1}(\xi)\|_{\infty} \geq O(2^{m-1}).
\]

IV. NUMERICAL EXPERIMENT

In this section, we provide some simple numerical simulations to verify some theoretical accuracy estimations in section III.

A. Experiment Setup

Suppose our filter \( a \) is supported in \( \{-2, \cdots, 2\} \). For example, \( \hat{a}(\xi) = 0.1 + 0.8 \cos(2\pi \xi) + 0.1 \cos(4\pi \xi) \), \( x \) is dirac at the center so that \( \hat{x}(\xi) = 1 \), and \( m = 3 \).

1) Choose \( \xi_1, \xi_2, \cdots, \xi_d \), and calculate \( \hat{y}_l(\xi_i) \) and the perturbed \( \hat{y}_l(\xi_i) = \hat{y}_l(\xi_i) + \epsilon_l \) for \( l = 0, \cdots, 5 \), where \( \epsilon_l \) is defined as in (1) and \( \epsilon_l \ll 1 \).

2) Use Algorithm II.1 to calculate the roots of \( p^\xi \) and the perturbed roots of \( \tilde{p}^\xi \) respectively, then compute \( |\Delta_k(\xi)| = |\hat{a}(\xi_i^m) - \tilde{a}(\xi_i^m)| \).

B. Experiment Results

1) Sharpness of estimation (6) and (8). Fix \( m \) and \( x \), our estimation (6) and (8) suggest that the error \( |\Delta_k(\xi)| = |
\hat{a}(\xi_i^m) - \tilde{a}(\xi_i^m)| \) in the worst possible case could be proportional to \( \delta_k(\xi)\delta(\xi) \). In this experiment, we choose six points \( \xi_1 < \cdots < \xi_6 < \frac{1}{2} \) such that \( \delta_0(\xi_i)\delta(\xi_i) \) grows geometrically at rate \( 10^3 \), and set the noise level \( \epsilon \sim 10^{-14} \). In Figure 1(a), we plot the value of \( \Delta_0(\xi_i) \) for \( i = 1, \cdots, 6 \). We can see that \( \Delta_0(\xi_i) \) grows proportionally to the growth of \( \delta_0(\xi_i)\delta(\xi_i) \), which verified the sharpness of estimation(6) and (8).

2) Dependence of \( \Delta_k(\xi) \) on the measurement error \( \epsilon \): In this experiment, we fix some \( 0 < \xi < \frac{1}{2} \), we only change the magnitude of the error \( \epsilon = \max \epsilon_l \) such \( \epsilon_l \ll 1 \) and that \( \epsilon \) grows geometrically at rate \( \sqrt{\epsilon} \) and keep other parameters unchanged. We plot the value of \( \Delta_k(\xi) \) for \( k = 1, \cdots, 5 \) in different noise levels. The results are presented in Figure 1(b). We show that \( |\Delta_k(\xi)| \sim O(\epsilon) \) for \( k = 1, 2 \).

3) The infinity norm of \( H_{m^{-1}}^{-1}(\xi) \): In this experiment, we choose \( m = 2, 3, \cdots, 6 \) and \( \xi = 0.3 \). We compute and plot the value of \( \|H_{m^{-1}}^{-1}(\xi)\|_{\infty} \) for different \( m \). The results are presented in Figure 1(c). It is shown that \( \|H_{m^{-1}}^{-1}(\xi)\|_{\infty} \) grows geometrically.

V. CONCLUSION

In this paper, we investigate the conditions under which we can recover a typical low pass convolution filter \( \hat{a} \in \ell^2(\mathbb{Z}) \) and a vector \( x \in \ell^2(\mathbb{Z}) \) from the combined regular subsampled version of the vector \( x, \cdots, A^{-1}x \) defined in (1), where \( A \) is a generalized Prony method is proposed to show that \( \{x|_{\Omega_m}, x(x(1)|_{\Omega_m}, \cdots, x(N)|_{\Omega_m} : N \geq 2m - 1, \Omega_m = m\mathbb{Z}\} \) contains enough information to recover \( a \) almost surely. Our accuracy estimates are formulated in very simple geometric terms involving Fourier spectral function of \( a, x \) and \( m \), shedding some light on the structure of the problem. Our results suggest that when the generalized Prony method is used, the parameters of the problem are coupled to each other, in the sense that the accuracy of recovering the nodes \( \{\hat{a}(\xi_i^m) : i = 0, \cdots, m - 1\} \) depends on the values of all the parameters at once. This unfavorable behavior is reflected by our numerical experiments in section IV. For example, the accuracy of recovering the node \( \hat{a}(\xi_i^m) \) not only depends on its separation with other nodes \( \delta_k(\xi) \) (see Defintion 2), but also...
depends on the minimum separation $\delta(\xi) = \max_{k=0, \ldots, m-1} \delta_k(\xi)$ among the nodes. The classical Prony method performs poorly when noisy sampled data are given. In our case, we have similar issues, since our Hankel matrix $H_m(\xi)$ is ill conditioned. In practice, we can employ denoising techniques to process sampled data such as Cadzow denoising algorithm to make the method more robust to noise. However, we believe that a full answer to our somewhat rigid $\ell^\infty$ formulation of the accuracy problem may contribute to the understanding of limitations of using Prony type methods in spatiotemporal sampling.

VI. FUTURE WORK
Finding stable solution of Prony-type systems is generally considered to be a difficult problem, and in recent years many algorithms have been devised for this task, such as nonlinear least squares, ESPRIT. We believe these algorithms can be adapted to our case and make our solution become more robust to noise, which we posit as our future work.

ACKNOWLEDGMENT
The author would like to thank Akram Aldroubi for his endless encouragements and insightful comments in the process of creating this work. The author also thanks Liyra Kristal and Yang Wang for their helpful discussions related to this work. The author is indebted to anonymous referees for their useful comments to improve the quality of the presentation of this paper. This research was supported in part by NSF Grant DMS-1322099.

REFERENCES