RECONSTRUCTION OF THE SEQUENCE OF DIRACS FROM NOISY SAMPLES VIA MAXIMUM LIKELIHOOD ESTIMATION

Akira Hirabayashi[1] , *Takuya Iwami*[1], *Shuji Maeda*[2], *and Yosuke Hironaga*[2]

[1]Graduate School of Medicine, Yamaguchi University
[2]Faculty of Engineering, Yamaguchi University
2-16-1, Tokiwadai, Ube, Yamaguchi 755-8611, Japan

ABSTRACT

We propose a reconstruction procedure for periodic sequence of K Diracs from noisy uniform measurements based on the maximum likelihood estimation. We first express the noise vector using the measurement vector and estimation parameters. This expression and the probability density function (PDF) for the noise vector allow us to define the (log-)likelihood function. We show that when the PDF is Gaussian, the maximization of the likelihood function is equivalent to finding the nearest sequence to the noisy sequence in the Fourier domain. This problem can be efficiently solved by combining an analytic solution and the so-called particle swarm optimization (PSO) search. Computer simulations show that the proposed method outperforms the conventional methods with computational cost of approximately $O(K)$.

Index Terms— Sequence of Diracs, signals with finite rate of innovation, annihilating filter, maximum likelihood estimation

1. INTRODUCTION

One of the recent hot topics in sampling theory is sampling for signals with a finite rate of innovation (FRI) [1], [2], [3]. A signal in this class has a parametric representation, and the rate of innovation, which is defined by the number of parameters in an arbitrary interval divided by the length of the interval, is finite. For example, the sequence of Diracs, non-uniform polynomial splines, piecewise polynomials, and piecewise sinusoids [4] are included in the class.

Let us focus on the periodic sequence of K Diracs in this paper. The representation parameters of this signal are the location of Diracs and their intensity coefficients. The key feature of the signal is that the Fourier coefficients are given by a linear combination of exponentials. Each of the exponentials can be identified from the Fourier coefficients by using the so-called annihilating filter. That is, we determine an auto regressive filter that transforms the Fourier coefficients to a zero sequence. Factorization of the filter coefficients provides each exponent, and then the location of Diracs is determined. The intensity coefficients can be easily found thereafter.

When signals are degraded by noise, annihilating equations do not yield zero any more. The relevant solution to this problem is to minimize the sum of squared residues of annihilating equations [3]. Maravic *et al.* proposed the so-called subspace-based approach [5]. Neither of these methods, however, exploits statistical information about noise. Hence, in this paper, we propose a method that uses the information based on the maximum likelihood estimation. We first express noise vector by noisy measurement vector and estimation parameters (locations of Diracs and intensity coefficients). Substituting the expression to the probability density function (PDF) for the noise vector leads us to the (log-)likelihood function. We then assume that the PDF is Gaussian and derive an explicit formula for the log-likelihood function, which shows that the maximum likelihood estimation is equivalent to finding the nearest sequence to the noisy sequence in the Fourier domain. This problem can be efficiently solved by combining an analytic solution and the so-called particle swarm optimization (PSO) search [6]. Computer simulations show that the proposed method outperforms the least-squares and subspace-based approaches for all signal-to-noise ratios from 0 to 50 [dB] with reasonable computational costs of approximately $O(K)$.

This paper is organized as follows. In Section 2 we summarize the sampling and reconstruction of the periodic sequence of Diracs from noiseless measurements. In Section 3, we take noise in measurements into account and define the likelihood or log-likelihood functions. In Section 4, we derive an explicit formula of the log-likelihood function for a Gaussian distribution, and propose a method to find the best solution to the criterion. Computer simulations are also performed in this section. Section 5 concludes the paper.

2. NOISELESS FORMULATION

Assume that we observe a τ-periodic sequence of $K(<\infty)$ Diracs, $s(t)$, which is expressed using its single period signal

This work was partially supported by KAKENHI, Grant-in-Aid for Scientific Research (C), 23500212, 2011.

$s_0(t)$ in the interval $[0, \tau)$ as

$$s(t) = \sum_{k'=-\infty}^{\infty} s_0(t - k'\tau),$$

where

$$s_0(t) = \sum_{k=0}^{K-1} c_k \delta(t - t_k),$$

and where $0 \leq t_0 < t_1 < \ldots < t_{K-1} < \tau$. The rate of innovation is $\rho = 2K/\tau$, which is finite because K is finite. Since $s(t)$ is periodic, it can be expressed by the Fourier series, in which the coefficient is given by

$$\hat{d}_p = \frac{1}{\tau}\int_0^{\tau} s(t)e^{-i2p\pi t/\tau}dt = \frac{1}{\tau}\sum_{k=0}^{K-1} c_k u_k^p, \tag{1}$$

where $u_k = e^{-i2\pi t_k/\tau}$.

Noiseless measurements $\{d_n\}_{n=0}^{N-1}$ are given by

$$d_n = \langle s, \psi_n \rangle = \int_{-\infty}^{\infty} f(t)\overline{\psi(t - nT)}dt, \tag{2}$$

where $T = \tau/N$. The sampling kernel is $\psi(t) = B\mathrm{sinc}(Bt)$, where $B \geq \rho = 2K/\tau$ and $\mathrm{sinc}(t)$ is $\sin(\pi t)/(\pi t)$ if $t \neq 0$, or 1 if $t = 0$.

The unknown parameters $\{t_k\}_{k=0}^{K-1}$ and $\{c_k\}_{k=0}^{K-1}$ can be retrieved from the measurements $\{d_n\}_{n=0}^{N-1}$ by the so-called annihilating filter [1]. Let the filter coefficients be $[1, a_1, \ldots, a_K]$. Then, the annihilating equations are

$$\hat{d}_p + a_1\hat{d}_{p-1} + \ldots + a_K\hat{d}_{p-K} = 0 \tag{3}$$

for $p = K - P, \ldots, P$, where $P = \lfloor B\tau/2 \rfloor$, the maximum integer not exceeding $B\tau/2$. The Fourier coefficients $\hat{d}_p$ are obtained by the discrete Fourier transform (DFT) from the measurements d_n because of the sinc sampling. Simultaneously solving the annihilating equations yields the filter coefficients, and factorization provides u_k, and thus t_k. We need $2K$ values of $\hat{d}_p$ in the K equations of (3). Because of symmetry of the Fourier series to express real numbers d_n, the number of coefficients $\hat{d}_p$ must be odd. Hence, we need at least $2K + 1$ measurements of d_n: $N \geq 2P + 1$. Once t_k or u_k is obtained, we can determine c_k by (1), which completes the reconstruction for the noiseless case. For simplicity, we consider the case where $N = 2P + 1$.

Let $\boldsymbol{d}$ and $\hat{\boldsymbol{d}}$ be the N dimensional vectors whose n-th elements are d_n and $\hat{d}_{n-P}$, respectively. Then, it holds that

$$\hat{\boldsymbol{d}} = F\boldsymbol{d}, \tag{4}$$

where F is the DFT matrix. Let $\boldsymbol{t}$ and $\boldsymbol{c}$ be K dimensional vector whose k-th elements are t_k and c_k, respectively. Further, let us define an $N \times K$ matrix U_t by

$$U_t = \begin{pmatrix} u_0^{-P} & u_1^{-P} & \ldots & u_{K-1}^{-P} \\ u_0^{-P+1} & u_1^{-P+1} & \ldots & u_{K-1}^{-P+1} \\ \vdots & \vdots & \ddots & \vdots \\ u_0^{P} & u_1^{P} & \ldots & u_{K-1}^{P} \end{pmatrix}.$$

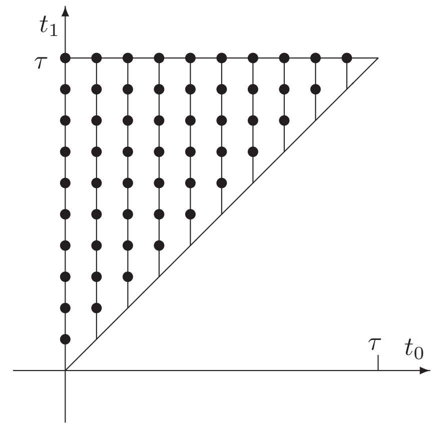

Fig. 1. Search space for the optimization of $f(t_0, t_1)$. The coarse search was conducted on the dotted pairs of (t_0, t_1). The fine search was done around the optimal pair over twenty times finer grid.

Then, (1) yields the matrix expression as

$$\hat{\boldsymbol{d}} = U_t\boldsymbol{c}. \tag{5}$$

3. ML ESTIMATION FROM NOISY MEASUREMENTS

Now, we consider the case in which measurements d_n are corrupted by noise e_n as $y_n = d_n + e_n$. Let $\boldsymbol{y}$ and $\boldsymbol{e}$ stand for the vectors whose n-th elements are y_n and e_n, respectively. Then, we have

$$\boldsymbol{y} = \boldsymbol{d} + \boldsymbol{e}. \tag{6}$$

In this case, the corresponding annihilating equation does not hold for any coefficients a_k in general:

$$\hat{y}_p + a_1\hat{y}_{p-1} + \ldots + a_K\hat{y}_{p-K} \neq 0,$$

where $\hat{y}_p$ is the $(p+P)$-th element of the vector $\hat{\boldsymbol{y}} = F\boldsymbol{y}$. The relevant solution would be to minimize the sum of squares of the left-hand side [3]. However, this does not always provide a good solution because the noise level might not be moderate. To obtain stable results even when noise level is severe, we propose the use of maximum likelihood (ML) estimation.

To this end, we define a likelihood function as follows. From (6), (4), and (5), we have

$$\boldsymbol{e} = \boldsymbol{y} - \boldsymbol{d} = \boldsymbol{y} - F^{-1}\hat{\boldsymbol{d}} = \boldsymbol{y} - F^{-1}U_t\boldsymbol{c}. \tag{7}$$

Now, the noise vector $\boldsymbol{e}$ is expressed using the sample vector $\boldsymbol{y}$ and the estimation parameter vectors $\boldsymbol{t}$ and $\boldsymbol{c}$.

We assume that the probability density function $p(\boldsymbol{e})$ is known. Substituting (7) into $p(\boldsymbol{e})$ enables us to define the likelihood function

$$L(\boldsymbol{t}, \boldsymbol{c}) = p(\boldsymbol{y} - F^{-1}U_t\boldsymbol{c}),$$

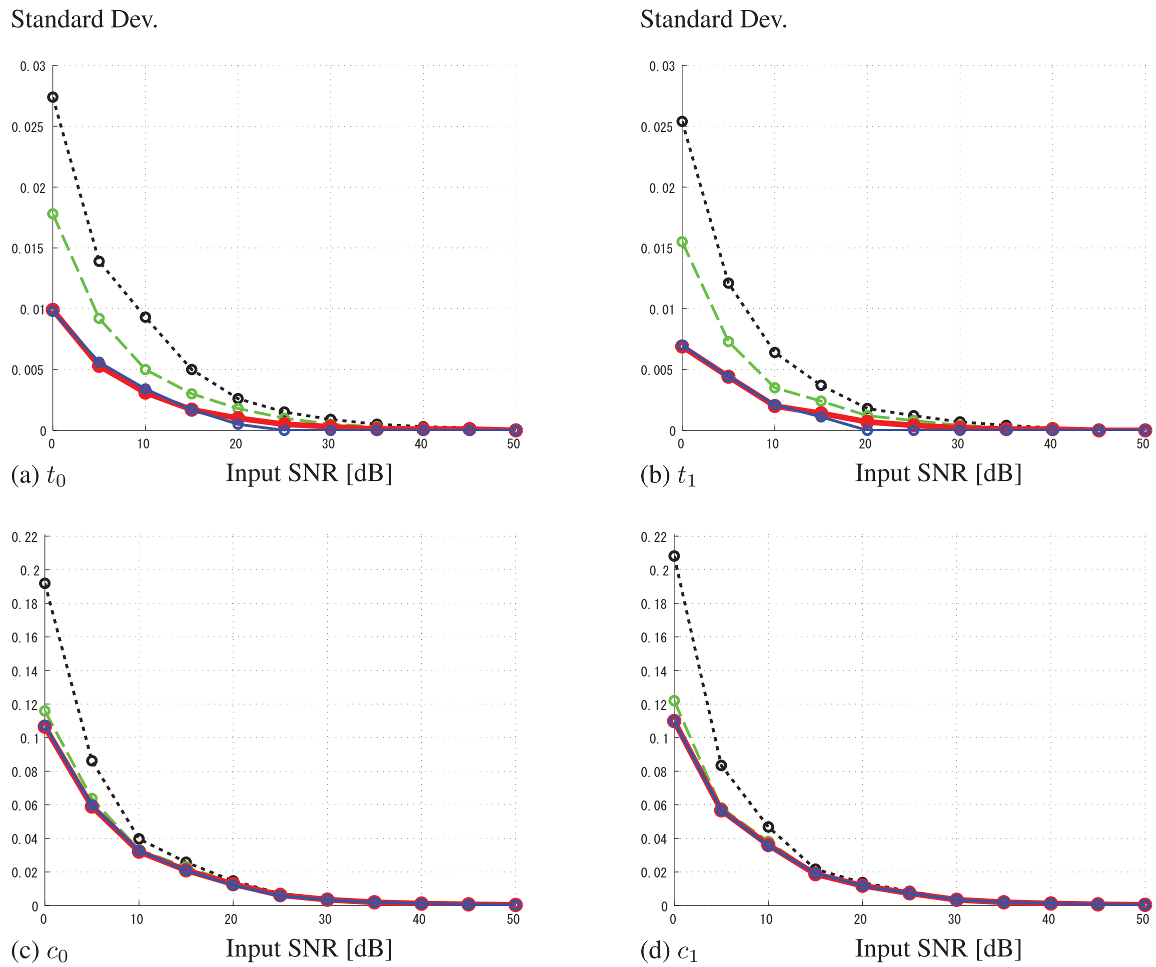

Fig. 2. Standard deviations of the estimated values for (a) t_0, (b) t_1, (c) c_0, and (d) c_1 when the input signal-to-noise ratio was 0[dB] $\sim$ 50[dB]. The red/thick and blue/thin solid lines show the results by the proposed methods with particle swarm optimization and coarse-to-fine search, respectively. The black dotted and green dashed lines show the results by the least-squares method and the subspace-based approach, respectively.

or the log-likelihood function

$$l(\boldsymbol{t}, \boldsymbol{c}) = \log p(\boldsymbol{y} - F^{-1} U_t \boldsymbol{c}).$$

By maximizing either of these functions, we can precisely estimate $\{t_k\}_{k=0}^{K-1}$ and $\{c_k\}_{k=0}^{K-1}$.

4. GAUSSIAN NOISE

Assume that $p(\boldsymbol{e})$ is the Gaussian distribution with zero mean and the covariance matrix $\sigma^2 I$, where σ is a known positive number and I is the identity matrix. Then, the log-likelihood function yields

$$l(\boldsymbol{t}, \boldsymbol{c}) = -\frac{\|\boldsymbol{y} - F^{-1} U_t \boldsymbol{c}\|^2}{2\sigma^2} + \text{Const}.$$

This implies that the maximization of the log-likelihood function $l(\boldsymbol{t}, \boldsymbol{c})$ is equivalent to the minimization of the norm $\|\boldsymbol{y} - F^{-1} U_t \boldsymbol{c}\|^2$, which is further equivalent to that of

$$f_o(\boldsymbol{t}, \boldsymbol{c}) = \|\hat{\boldsymbol{y}} - U_t \boldsymbol{c}\|^2, \tag{8}$$

where $\hat{\boldsymbol{y}} = F\boldsymbol{y}$. Equation (8) implies that the maximum likelihood estimation finally resulted in the least mean square estimation in the Fourier domain. We now solve the minimization problem of (8). Note that (8) is quadratic in terms of $\boldsymbol{c}$. Hence, its optimal solution for fixed $\boldsymbol{t}$ can be determined as $\boldsymbol{c} = U_t^\dagger \hat{\boldsymbol{y}}$, where $U_t^\dagger$ is the Moore-Penrose generalized inverse of U_t. Hence, $\boldsymbol{t}$ and $\boldsymbol{c}$ that minimizes $f_o(\boldsymbol{t}, \boldsymbol{c})$ in (8) are given by $\boldsymbol{t}$ which minimizes

$$f(\boldsymbol{t}) = f_o(\boldsymbol{t}, U_t^\dagger \hat{\boldsymbol{y}}) = \|\hat{\boldsymbol{y}} - U_t U_t^\dagger \hat{\boldsymbol{y}}\|^2, \tag{9}$$

and $\boldsymbol{c} = U_t^\dagger \hat{\boldsymbol{y}}$ with the resultant $\boldsymbol{t}$. Further, because of the order relation among $\{t_k\}_{k=0}^{K-1}$, we can restrict the search space. For example, when $K = 2$, the two parameters t_0 and t_1 are looked for in the triangle area shown in Fig. 1.

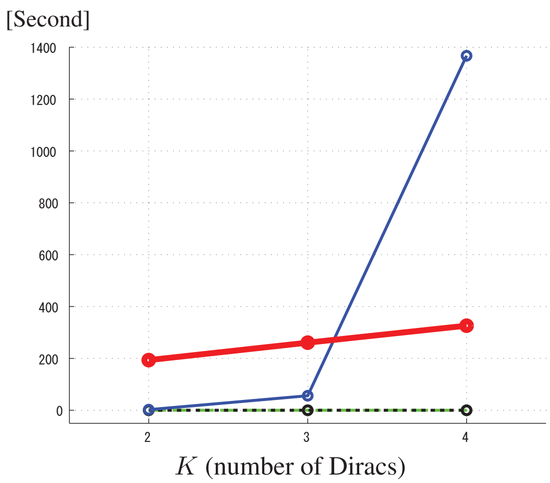

Fig. 3. Computational cost. The same legends are used as in Fig. 2.

To find the optimum parameters in the space, we adopted two optimization approach. One is a coarse-to-fine search. First, $f(t_0, t_1)$ is evaluated on the dotted pairs of (t_0, t_1), and then it is further evaluated over a 21×21 square grid around the optimal pair during the coarse search. The grid in the second search is twenty times finer than the coarse one.

The other is a particle swarm optimization (PSO) [6], which is a population based stochastic optimization technique. The algorithm is first initialized with a group of random particles and then searches for the optimal solution by updating particles by both the best solution for each particle achieved so far and the global best solution in the population. To avoid divergence of the algorithm, we suppressed particle movement as iteration increases. We used thirty particles and stopped computation after four hundred iterations.

We measured $K = 2$ Diracs located at $t_0 = 0.19$ and $t_1 = 0.63$. The period was $\tau = 1$. The coefficients were $c_0 = 0.8$ and $c_1 = 1.0$, respectively. The number of samples was $N = 9$ and then the redundancy was 2.25.

A hundred of noise vectors, $\boldsymbol{e}$, were generated from the Gaussian distribution in which σ was determined so that the signal-to-noise ratio (SNR) becomes $0, 5, \ldots, 50$ [dB]. We estimated t_0, t_1, c_0, and c_1 from the noisy measurements using the proposed method as well as the least-squares method [3] and subspace-based approach [5]. The standard deviations of the estimated values for t_0, t_1, c_0, and c_1 are shown in Fig. 2 (a), (b), (c), and (d), respectively. The red/thick and blue/thin solid lines show the results obtained by the proposed methods with PSO and coarse-to-fine search, respectively. The black dotted and green dashed lines show the results by the least-squares method and the subspace-based approach, respectively. We can see that the proposed method provides more stable results than those of the conventional methods. Since all of the methods produce the true values from noiseless measurements, they perform mostly same when SNR is high. When SNR is low, however, the proposed methods outperform both of the conventional methods. For example, when SNR is 0 [dB], the proposed method improved the standard deviation by 44% and 8% for t_0 and c_0, respectively. The same tendency was seen for t_1 and c_1 as well.

Computational costs were compared in Fig. 3. We can see that the proposed methods need more computation than the conventional ones. It is still interesting to note that the computational cost for the PSO method increases almost linearly while the coarse-to-fine search needs exponential cost. To sum up, the proposed method with PSO is capable of providing stable results with reasonable computational cost.

5. CONCLUSION

We proposed a reconstruction procedure for the periodic sequence of K Diracs from noisy uniform measurements based on the maximum likelihood estimation. We showed that when the probability density function is Gaussian, the maximization of the likelihood function is equivalent to finding the nearest sequence to the noisy sequence in the Fourier domain. This problem was efficiently solved by the particle swarm optimization search. Computer simulations showed that the proposed method outperformed the conventional least-squares approach for all SNR from 0[dB] to 50[dB] with computational costs of approximately $O(K)$.

6. REFERENCES

[1] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Trans. Signal Processing*, vol. 50, no. 6, pp. 1417–1428, June 2002.

[2] P.L. Dragotti, M. Vetterli, and T. Blu, "Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang-Fix," *IEEE Trans. Signal Processing*, vol. 55, no. 5, pp. 1741–1757, May 2007.

[3] T. Blu, P.L. Dragotti, M. Vetterli, P. Marziliano, and L. Coulot, "Sparse sampling of signal innovations," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 31–40, March 2008.

[4] J. Berent, P.L. Dragotti, and T. Blu, "Sampling piecewise sinusoidal signals with finite rate of innovation methods," *IEEE Trans. Signal Processing*, vol. 58, no. 2, pp. 613–625, February 2010.

[5] I. Maravic and M. Vetterli, "Sampling and reconstruction of signals with finite rate of innovation in the presence of noise," *IEEE Trans. Signal Processing*, vol. 53, no. 8, pp. 2788–2805, August 2005.

[6] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. International Conference on Neural Networks (ICNN 1995)*, Perth, 1995, vol. 4, pp. 1942-1948.