H^∞ Optimal Nonparametric Density Estimation from Quantized Samples

M. Nagahara[1], K. I. Sato[2], and Y. Yamamoto[1]
[1]Graduate School of Informatics, Kyoto University,
[2]Graduate School of Economics, Kyoto University,
Sakyo-ku Yoshida-Honmachi, Kyoto, 606-8501, Japan

Abstract

In this paper, we study nonparametric density estimation from quantized samples. Since quantization decreases the amount of information, interpolation (or estimation) of the missing information is needed. To achieve this, we introduce sampled-data H^∞ control theory to optimize the worst case error between the original probability density function and the estimation. The optimization is formulated by linear matrix inequalities and equalities. A numerical example is illustrated to show the effectiveness of our method.

1 Introduction

Probability density functions (PDF's, henceforth) or probability distributions play a fundamental role in both analysis and design of stochastic systems. Theoretically, it is often assumed that such information about randomness is given a priori. However, in practice, we usually have only limited information about that. On one hand, the random nature of such systems hinders us to obtain the exact distribution. On the other hand, because the analysis and design are usually executed on a digital processor, acquired data must be discretized in both time and value and this severely reduces the information. In this paper, we propose a method that weakens the latter difficulty. That is, we are concerned about estimating a PDF from quantized data. Our analysis assumes no specific function form for the PDF, and hence our method is supposed to be classified into nonparametric estimation. At the same time, it might be worth mentioning that we exploit a couple of ideas which were given rise in the signal processing and control literatures.

The first idea that we rely on owes Widrow [13, 14]. These papers study the estimation problem from quantized samples by applying Shannon's sampling theorem [8, 10]. We refer to the Widrow's result as Widrow's quantizetion theorem. This theorem assumes that the original characteristic function is band-limited up to the frequency π/Δ, where Δ is the step size of the uniform quantizer. This frequency corresponds to the Nyquist frequency in Shannon's sampling theorem. The assumption which is called *band-limiting condition* is quite restrictive because many random variables such as uniform, Gaussian, Erlangian, or Rayleigh ones do not satisfy the assumption [6].

The second helpful idea has come from the control theory. In order to avoid the stringent assumption of Shannon's theorem, sampled-data H^∞ infty control theory proved useful in [5]. The paper shows that the sampled-data H^∞ optimization provides a way to obtain a filter without the band-limiting assumption. By altering the method of [5] into that suitable for Widrow's quantization theorem, we will be able to obtain a better estimation.

In this paper, this alternative way to estimating a PDF is proposed. Unlike Widrow's method, we are unable to obtain the exact PDF. Such a restriction is inevitable because we dismiss the band-limiting condition. Instead, we apply the sampled-data control theory [2] that enables us to minimize the undesirable error in the analog domain. A designed filter, which generates an estimated PDF, will be combined with an up-sampler to compose an interpolator [11]. Like the method of [5], our estimation can be accomplished by efficient numerical optimization.

The organization of this paper is as follows. Section 2 defines nonparametric density estimation problem which we discuss in this paper. In section 3, we introduce Widrow's quantization theorem, and we point out problems in Widrow's estimation. To solve these problems, we propose a new estimation method by sampled-data H^∞ optimization in section 4. Numerical examples, given in section 5, illustrate the effectiveness of our method compared with Widrow's estimation. Section 6 concludes our results and shows a future work.

Notations

We use the following notations. $\mathbb{R}$ and $\mathbb{Z}$ are the sets of the real numbers and the integers, respectively. L^2 is the Lebesgue spaces consisting of the square integrable real functions on $\mathbb{R}$. For $w \in L^2$, $\mathcal{F}w$ denotes the Fourier transform of w. We denote by χ_I the indicator function on a set $I \subset \mathbb{R}$; $\chi_I(x) = 1$ if $x \in I$ and $\chi_I(x) = 0$ if $x \notin I$. For a matrix M, $M^\top$ and tr$M$ are respectively the transpose and the trace of M. For a symmetric matrix P, $P \succ 0$, $P \succeq 0$ and $P \prec 0$ denote positive, nonnegative, and respectively negative definite matrices. For a linear time-invariant (continuous-time or discrete-time) system with state space matrices

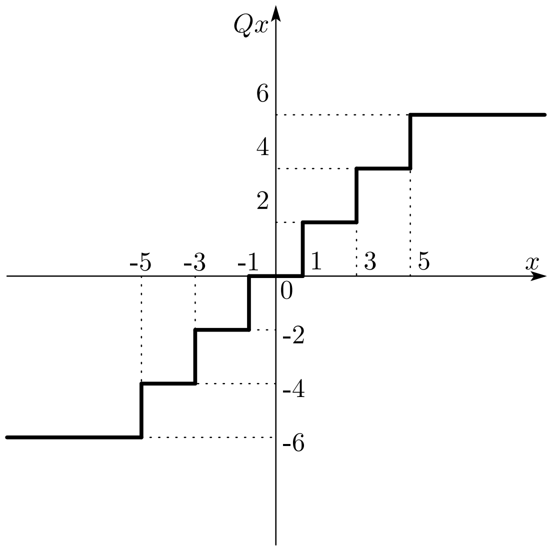

Fig. 1: Uniform quantizer with $\Delta = 2$ and $M = 3$

$\{A, B, C, D\}$,

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right](s) \quad \text{or} \quad \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right](z)$$

denotes the transfer function.

2 Density Estimation Problem

In this section, we define our problem. Assume that there are samples drawn from a PDF $p(x)$, which is unknown. Our concern in the present paper is how to estimate $p(x)$ from the samples. In ordinary practice the acquired samples are stored and processed in digital systems, and hence they must be converted to digital data, that is quantized. We assume that this quantization is a uniform one with the step size Δ and the no-overload input range $((-M-1)\Delta, (M+1)\Delta)$ (see [7]), which is defined by

$$Q : \mathbb{R} \to \mathcal{A} := \{-M\Delta, (-M+1)\Delta, \ldots, M\Delta\},$$
$$Qx := \begin{cases} k\Delta, & \left(k - \frac{1}{2}\right)\Delta \le x < \left(k + \frac{1}{2}\right)\Delta, \\ & \quad k = -M+1, \ldots, M-1, \\ M\Delta, & x \ge \left(M - \frac{1}{2}\right)\Delta, \\ -M\Delta, & x < \left(-M - \frac{1}{2}\right)\Delta. \end{cases} \tag{1}$$

Fig. 1 shows a uniform quantizer with $\Delta = 2$ and $M = 3$.

Then, our problem is formulated as follows:

Problem 1 *Given samples $\widetilde{x}_1, \widetilde{x}_2, \ldots, \widetilde{x}_N$ which are i.i.d. (independent, identically distributed) and quantized by the uniform quantizer defined by (1), estimate the original PDF $p(x)$ from which the samples are drawn.*

This problem is ill-posed because there are infinitely many solutions. To avoid ill-posedness, we have to restrict the function space to which the original PDF belongs. A unique solution to the problem is obtained by Widrow et al. by restricting the function space to that of band-limited functions [13, 14]. We introduce this solution in the next section.

3 Widrow's Quantization Theorem

In this section, we introduce Widrow's quantization theorem.

First, we analyze the uniform quantization Q (see (1)) of a PDF $p(x)$. The samples drawn from $p(x)$ and quantized by Q can be interpreted as samples of a discrete PDF $c[k]$. This discrete PDF is obtained by *area-sampling* [13, 14], denoted by $\widetilde{\mathcal{S}}_\Delta p$, that is,

$$c[k] = \int_{(k-1/2)\Delta}^{(k+1/2)\Delta} p(x)dx =: (\widetilde{\mathcal{S}}_\Delta p)[k] \tag{2}$$

This can be rewritten by

$$\begin{aligned} c[k] &= \int_{-\infty}^{\infty} \chi_{[-\Delta/2,\Delta/2)}(k\Delta - x)p(x)dx \\ &= (q_\Delta * p)(k\Delta) =: \mathcal{S}_\Delta(q_\Delta * p)[k], \\ q_\Delta(x) &:= \chi_{[-\Delta/2,\Delta/2)}(x). \end{aligned}$$

Thus, quantization can be interpreted as convolution by q_Δ and ideal sampling $\mathcal{S}_\Delta$. Note that the Fourier transform (frequency response) of q_Δ is given by

$$Q_\Delta(j\omega) = \frac{1}{j\omega}\left(e^{j\omega\frac{2}{\Delta}} - e^{-j\omega\frac{2}{\Delta}}\right). \tag{3}$$

Based on this idea, Widrow et al. proposed a reconstruction scheme by using Shannon's sampling theorem. To see this, we define the space BL^2 of band-limited functions:

$$BL^2 := \left\{p \in L^2 : (\mathcal{F}p)(\omega) = 0, |\omega| > \pi/\Delta\right\}.$$

Then, Widrow et al. have shown a solution to Problem 1 as follows [13, 14].

Theorem 1 (Widrow et al.) *Assume that $p \in BL^2$. Then $p(x)$ can be perfectly reconstructed from the discrete PDF of the quantized samples $\{\widetilde{x}_n\}$.*

The reconstruction procedure by this theorem is as follows. First, construct a histogram from the samples $\{\widetilde{x}_n\}$, with its intervals $(-\infty, -(M+1/2)\Delta)$, $[(k-1/2)\Delta, (k+1/2)\Delta)$, $k = -M+1, \ldots, M-1$, and $[(M-1/2)\Delta, \infty)$. If we have sufficiently many samples, $\{c_k\}$ can be approximated from the frequency of each interval divided by the number of the samples. Let $r = q_\Delta * p$. Then, if the band-limiting assumption holds, r can be perfectly reconstructed from the discrete data $\{c[k]\}$ by

$$r(x) = \sum_{k=-\infty}^{\infty} c[k]\mathrm{sinc}(x/\Delta - k) =: (\widetilde{\mathcal{H}}_{\mathrm{sinc}}c)(x),$$

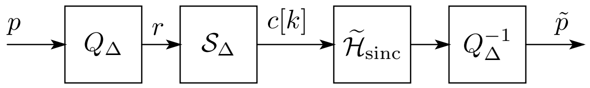

Fig. 2: Density estimation by Widrow's method

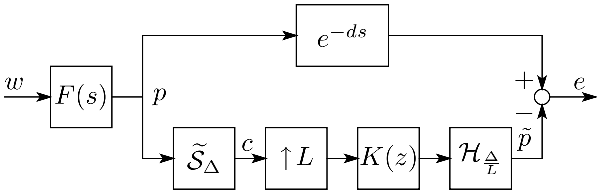

Fig. 3: Error system $\mathcal{E}$

where sinc is the sinc function defined by

$$\mathrm{sinc}(x) := \frac{\sin \pi x}{\pi x}.$$

Then, the original PDF $p(x)$ can be obtained by the convolution $p = r * q_\Delta^{-1}$, where q_Δ^{-1} is the inverse system of q_Δ, whose Fourier transform is given by (see (3))

$$Q_\Delta^{-1}(j\omega) = \frac{j\omega}{e^{j\omega\frac{2}{\Delta}} - e^{-j\omega\frac{2}{\Delta}}}.$$

This procedure is shown in Fig. 2.

4 H^∞ Optimal Estimation

4.1 Signal space FL^2

In practice, the band-limiting assumption $p \in BL^2$ can be too restrictive because many important probablity density functions such as uniform, Erlangian, or Rayleigh ones do *not* satisfy the assumption (see [6]). We therefore introduce another signal space FL^2 defined by

$$FL^2 := \{p \in L^2 : p = Fw, w \in L^2\},$$

where F is a stable and strictly causal system with a rational transfer function $F(s)$. Although this space does not include band-limited functions (since the sinc function does not have a rational transfer function), this space includes many functions such as the distributions mentioned above [6].

The function $F(s)$ is treated as a *smoothing parameter* [1] in our density estimation. For selecting $F(s)$, see sub-section 4.5 and section 5.

4.2 Density estimation via sampled-data H^∞ optimization

Assuming that the original PDF $p(x)$ is in FL^2, we consider the reconstruction problem by the error system shown in Fig. 3. In this figure, an L^2 signal w is filtered by $F(s)$. This $F(s)$ is a presumed frequency domain model for the original PDF p. Then, this p is sampled by the generalized sampler $\widetilde{\mathcal{S}}_\Delta$ defined by (2) and a discrete PDF $\{c[k]\}$ is obtained. To reconstruct p from the discrete data $\{c[k]\}$, we adopt the interpolation technique in multirate signal processing [11]. First, the discrete data $\{c[k]\}$ is upsampled by $\uparrow L$, which is defined by

$$((\uparrow L)c)[n] = \begin{cases} c[k], & n = Lk,\ k = 0, 1, \dots \\ 0, & \text{otherwise} \end{cases}$$

Then the upsampled signal $(\uparrow L)c$ is filtered by $K(z)$ which we design. The output of this filter is converted to a continuous function by the zero-order hold $\mathcal{H}_{\Delta/L}$ with period Δ/L, and we obtain an estimation $\tilde{p}$ of the original PDF p. In the error system Fig. 3, the estimation $\tilde{p}$ is compared with d-shifted version $p(x-d)$ of the original PDF $p(x)$. The objective is to attenuate the *continuous* reconstruction error e with respect to the H^∞ norm of the error system

$$\mathcal{E} := \{e^{-ds} - \mathcal{H}_{\Delta/L} K(z)(\uparrow L)\widetilde{\mathcal{S}}_\Delta\} F(s). \quad (4)$$

The optimal filter $K(z)$ is then obtained by modern sampled-data control theory, and has the following advantages:

1. the optimal filter is always stable;
2. the design takes the inter-sample behavior into account;
3. the system is robust against the uncertainty of F;
4. the optimal FIR (finite impulse response) filter is also obtainable via Linear Matrix Inequalities (LMI) assuming that the filter $K(z)$ is an FIR filter.

The last advantage is the key idea of our density estimation to consider constraints for PDF estimation. We discuss this in the next section.

4.3 Constraints for PDF estimation

Since $p(x)$ is a PDF, it satisfies

$$p(x) \geq 0, \quad (5)$$

$$\int_{-\infty}^{\infty} p(x)dx = 1. \quad (6)$$

We therefore make the estimated PDF $\tilde{p}(x)$ so as to meet the constraints (5) and (6). We here assume that the filter $K(z)$ is an FIR filter, that is,

$$K(z) = \sum_{n=0}^{N} a_n z^{-n}. \quad (7)$$

For mathematical reason, we define $a_n = 0$ for $n > N$ or $n < 0$. Then, the constraints (5) and (6) can be described by an LMI and an LME (linear matrix equality).

Lemma 1 *The estimation $\tilde{p}(x) \geq 0$ for any $x \in \mathbb{R}$ if and only if*

$$a_n \geq 0, \quad n = 0, 1, \dots, N.$$

Proof

Let $\{d[n]\}$ be the output of the filter $K(z)$, that is, $d := K(\uparrow L)c$. Then $\tilde{p}(x)$ is the output of the zero-order hold $\mathcal{H}_{\Delta/L}$ for $d[n]$, and it is clear that $\tilde{p}(x) \geq 0$ for any $x \in \mathbb{R}$ if and only if $d[n] \geq 0$ for any $n \in \mathbb{Z}$. Therefore, we show $d[n] \geq 0$ for any $n \in \mathbb{Z}$ if and only if $a_n \geq 0$ for $n = 0, 1, \ldots, N$.

First assume $a_n \geq 0$, $n = 0, 1, \ldots, N$. Then, we have

$$d[n] = \{(\uparrow Lc) * a)[n] = \sum_{k=-\infty}^{\infty} c[k] a_{n-Lk}. \tag{8}$$

Since $a_n = 0$ for $n > N$ and $n < 0$, and since $c[k] \geq 0$ for all $k \in \mathbb{Z}$, we have $d[n] \geq 0$, for all $n \in \mathbb{Z}$.

Conversely, assume that there exists a number $l \in \mathbb{Z}$ such that $a_l < 0$. Consider a discrete PDF

$$c[k] = \begin{cases} 1, & k = 0, \\ 0, & k \neq 0. \end{cases} \tag{9}$$

Then we have

$$d[L+l] = \sum_{k=-\infty}^{\infty} c[k] a_{L+l-Lk} = c[0] a_l = a_l < 0.$$

$\square$

Lemma 2 *Assume that*

$$\Delta \sum_{k=-\infty}^{\infty} c[k] = 1.$$

Then,

$$\int_{-\infty}^{\infty} \tilde{p}(x) dx = 1 \tag{10}$$

if and only if

$$\sum_{n=0}^{N} a_n = L. \tag{11}$$

Proof

First, assume the equality (11). By using (8), we have

$$\begin{aligned}
\int_{-\infty}^{\infty} \tilde{p}(x) dx &= \frac{\Delta}{L} \sum_{n=-\infty}^{\infty} d[n] \\
&= \frac{\Delta}{L} \sum_{n=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} c[k] a_{n-Lk} \\
&= 1 - \Delta \sum_{k=-\infty}^{\infty} c[k] \\
&\quad + \frac{\Delta}{L} \sum_{n=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} c[k] a_{n-Lk} \\
&= 1 + \frac{\Delta}{L} \sum_{k=-\infty}^{\infty} c[k] \left\{ \sum_{n=-\infty}^{\infty} a_{n-Lk} - L \right\}.
\end{aligned}$$

Since $a_n = 0$ if $n > N$ or $n < 0$, we have

$$\sum_{n=0}^{N} a_n = L, \tag{12}$$

if and only if

$$\sum_{n=-\infty}^{\infty} a_{n-Lk} = L,$$

for all $k \in \mathbb{Z}$. Hence (10) holds.

Conversely, assume the equality (12) does not hold. Consider the discrete PDF $\{c[k]\}$ is given by (9). Then we have

$$1 = \int_{-\infty}^{\infty} \tilde{p}(x) dx = 1 + \frac{\Delta}{L} \left\{ \sum_{n=-\infty}^{\infty} a_n - L \right\}.$$

Since $\Delta > 0$ and $L > 0$, this leads to contradiction. $\square$

By these lemmas, we can design the filter $K(z)$ which satisfies the constraints (5) and (6). Note that these constraints are not satisfied in Widrow's estimation.

4.4 Optimal FIR filter design

In this section, we give a design formula for our density estimation. Let a state space realization of $F(s)$ be

$$F(s) = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array} \right] (s).$$

Then by (2), we have

$$\begin{aligned}
c[k] &= \int_{k\Delta}^{(k+1)\Delta} p\,(x + \Delta/2)\, dx \\
&= C \int_0^{\Delta} e^{A\theta} d\theta \int_0^{\infty} e^{A(k\Delta - \tau - \Delta/2)} B w(\tau) d\tau.
\end{aligned}$$

Define

$$F_\Delta(s) := \left[\begin{array}{c|c} A & B \\ \hline C_\Delta & 0 \end{array} \right] (s), \quad C_\Delta := C \int_0^{\Delta} e^{A\theta} d\theta,$$

and assume that the reconstruction delay d satisfies $d = (m + 1/2)\,\Delta$. Then we have a generalized plant representation for our H^∞ design shown in Fig. 4. In this figure, $e'(t) = e(t - \Delta/2)$, and the delay is an all-pass function, the H^∞ norm will be the same.

For the sampled-data system shown in Fig. 4 can be reduced to a discrete-time system via H^∞ discretization or fast-sample fast-hold (FSFH) method [2, 3, 5]. Assuming that the filter $K(z)$ is an FIR filter given in (7),we have the following discrete-time error system $E_d(z)$ for the sampled-data system (4) [3],

$$E_d(z) = T_1(z) + W(\boldsymbol{a}) T_2(z),$$

where $\boldsymbol{a} := [a_0, a_1, \ldots, a_{N-1}]$ is the vector of the FIR coefficients, and $W(\boldsymbol{a}) := \begin{bmatrix} a_0 I & a_1 I & \ldots & a_{N-1} I \end{bmatrix}$. Thanks to the $\boldsymbol{a}$-affine structure of $E_d(z)$, a state space representation of $E_d(z)$ is given by

$$E_d(z) = \left[\begin{array}{c|c} A_e & B_e \\ \hline C_e(\boldsymbol{a}) & D_e(\boldsymbol{a}) \end{array} \right] (z),$$

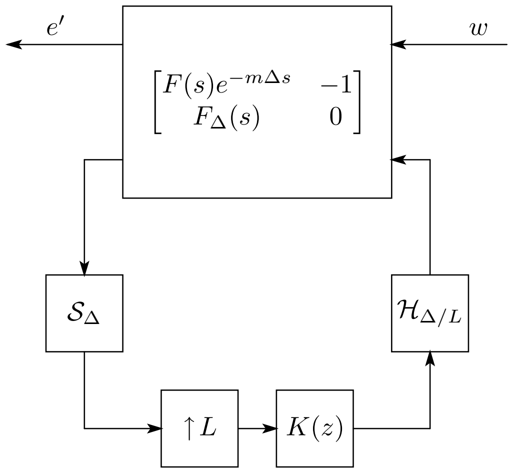

Fig. 4: Generalized plant for H^∞ optimal density estimation

where A_e and B_e are constant matrices and $C_e(\boldsymbol{a})$ and $D_e(\boldsymbol{a})$ affinely depend on the design parameter $\boldsymbol{a}$. By using the Kalman-Yakubovic-Popov lemma [4], our optimization can be described as follows. Find the optimal FIR coefficient $\boldsymbol{a}$ which minimizes $\gamma > 0$ subject to

$$P = P^\top \succ 0,$$
$$\begin{bmatrix} A_e^\top P A_e - P & A_e^\top P B_e & C_e(\boldsymbol{a})^\top \\ B_e^\top P A_e & -\gamma I + B_e^\top P B_e & D_e(\boldsymbol{a})^\top \\ C_e(\boldsymbol{a}) & D_e(\boldsymbol{a}) & -\gamma I \end{bmatrix} \prec 0,$$
$$V(\boldsymbol{a}) \succeq 0, \quad \mathrm{tr} V(\boldsymbol{a}) = L,$$

where $V(\boldsymbol{a}) = \mathrm{diag}\{a_1, a_2, \ldots, a_{N-1}\}$. The inequality $V(\boldsymbol{a}) \succeq 0$ is for the constraint (5) by Lemma 1, and the equality $\mathrm{tr}V(\boldsymbol{a}) = L$ is for (6) by Lemma 2. These are LMI's and an LME, and the optimal coefficient $\boldsymbol{a}$ can be obtained effectively by using an optimization softwares such as MATLAB. If the up-sampling ratio L is large (e.g., $L = 32$ or 64) and the number N of $K(z)$ is also large, the computation will be hard. In such a case, a cutting plane method can be applied. See [12].

4.5 Selection of parameters Δ and $F(s)$

We have assumed that the quantization step size Δ and the smoothing parameter $F(s)$ are given. However, in estimating a PDF, we have to select appropriate Δ and $F(s)$ for good estimation. An optimal method for selecting Δ is proposed in [9]. We can use this method for selecting Δ. On the other hand, $F(s)$ is typically chosen as a low-pass filter with a cutoff frequency less than the Nyquist frequency π/Δ. Although the cutoff frequency can be arbitrarily chosen in theory, the reconstructed PDF will be very cheap if the cutoff frequency is near or greater than π/Δ. So the cutoff frequency should be less than π/Δ. A method for selecting $F(s)$ is to estimate the cutoff frequency by interpolating and extrapolating the discrete Fourier transform of $\{c[k]\}$.

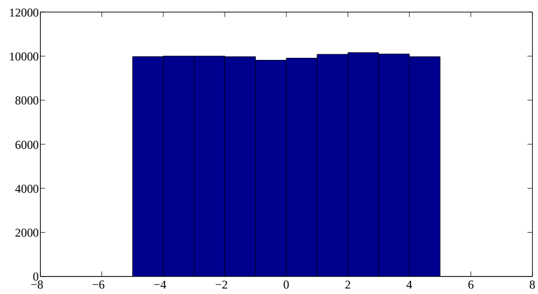

Fig. 5: Histogram of samples

5 Numerical Example

We here show an example of density estimation. We set the original PDF $p(x)$ to be the uniform distribution in the interval $[-5, 5]$. Note that since the support of $p(x)$ is $[-5, 5]$ (compact), $p(x)$ does not satisfy the band-limiting assumption of Widrow's theorem. The number of samples drawn from the PDF is 10^5. Then the samples are uniformly quantized with step size $\Delta = 1$. The histogram of the quantized samples is shown in Fig. 5.

For our design, we set the design parameters as follows. The transfer function (smoothing parameter) $F(s)$ is set as

$$F(s) = \frac{\omega_1 \omega_2}{(s+\omega_1)(s+\omega_2)}, \quad \omega_1 := \pi/10, \quad \omega_2 := \pi/20.$$

The up-sampling ratio L is 4. The number of taps is $N = 16$.

Fig. 6 shows the frequency response of H^∞ optimal filter $K(z)$, and Fig. 7 shows its impulse response. By Fig. 7 we can see that our filter satisfies the constraint $a_n \geq 0$ in Lemma 1. The constraint $\sum a_n = L$ in Lemma 2 means that $K(1) = L$ (i.e., the frequency response of $K(z)$ is L at $\omega = 0$), and Fig. 6 shows that this constraint is satisfied.

Fig. 8 shows the estimated PDFs by Widrow's method and the proposed one. It can be seen that the PDF of Widrow's estimation has negative values. On the other hand, our PDF is always positive. Moreover, our method well reconstructs the original PDF, while Widrow's estimation shows large ripples. The ripples are caused by the high frequency components in the original characteristic function $(\mathcal{F}p)(\omega)$. which is not considered in Widrow's theorem. By this example, we can see the efficiency of our method.

6 Conclusion

In this paper, we have proposed a new nonparametric density estimation from quantized samples via sampled-data H^∞ optimization. Our estimation is formulated by signal reconstruction with an up-sampler and a digital filter, which can be effectively used in digital systems

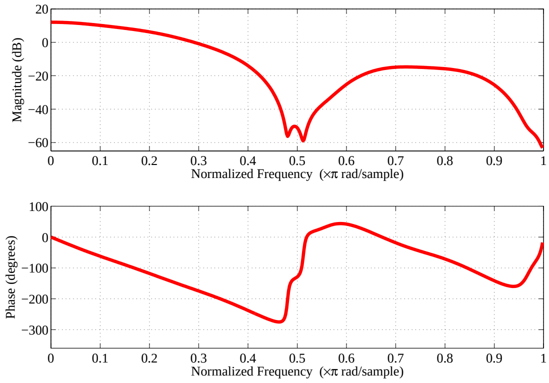

Fig. 6: Frequency response of H^∞ optimal filter $K(z)$

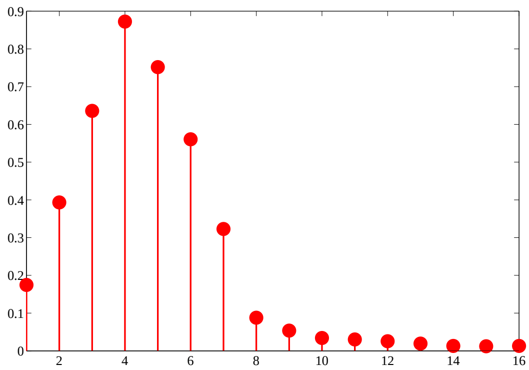

Fig. 7: FIR coefficients of H^∞ optimal filter $K(z)$

such as DSP's (Digital Signal Processors). The filter design is formulated by an H^∞ optimization, and the optimal filter can be computed by LMI's and an LME. We have studied density estimation of only one dimensional PDF. Future work includes higher dimensional case.

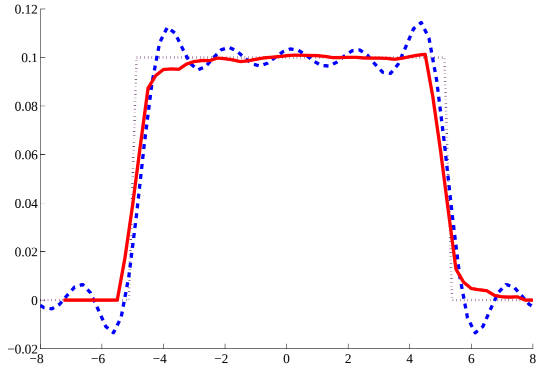

Fig. 8: Density estimation: original PDF (dots), Widrow's estimation (dash), and H^∞ optimal one (solid)

References

[1] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.

[2] T. Chen and B. A. Francis, *Optimal Sampled-data Control Systems,* Springer, 1995.

[3] H. Ishii and Y. Yamamoto, Sampled-data H^∞ and H^2/H^∞ design of multirate D/A converter, Trans. of the ISCIE, Vol. 11, No. 10, pp. 585–592, 1998 (in Japanese).

[4] T. Iwasaki, *LMI and Control*, Shoko-do, 1997 (in Japanese).

[5] M. Nagahara and Y. Yamamoto, A new design for sample-rate converters, *Proc. of 39th IEEE Conf. on Decision and Control*, pp. 4296–4301, 2000.

[6] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*, Fourth Ed., Mc Graw Hill, 2002.

[7] R. Schreier and G. C. Temes, *Understanding Delta-Sigma Data Converters*, Wiley Interscience, 2005.

[8] C. E. Shannon, Communication in the presence of noise, *Proc. IRE*, Vol. 37, No. 1, pp. 10–21, 1949.

[9] H. Shimazaki and S. Shinomoto, A method for selecting the bin size of a time histogram, *Neural Computation*, Vol. 19, pp. 1503-1527, 2007.

[10] M. Unser, Sampling — 50 years after Shannon, *Proc. IEEE*, Vol. 88, No. 4, pp. 569–587, 2000.

[11] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice Hall, Englewood Cliffs, 1993.

[12] Y. Wakasa, D. Yasufuku, M. Nagahara, and Y. Yamamoto, Sampled-data design of interpolators using the cutting-plane method, *Proc. of 41st IEEE Conf. on Decision and Control*, pp. 2564–2569, 2002.

[13] B. Widrow, A study of rough amplitude quantization by means of nyquist sampling theory, *IRE Trans. on Circuit Theory*, Vol. 3, No. 4, pp. 266 – 276, 1956.

[14] B. Widrow, I. Kollar, and M.-C. Liu, Statistical theory of quantization, *IEEE Trans. on Instrumentation and Measurement*, Vol. 45, No. 2, pp. 353 – 361, 1996.