

Source Coding for a Wiener Process

Sahan Liyanaarachchi Ismail Cosandal Sennur Ulukus
 Department of Electrical and Computer Engineering
 University of Maryland, College Park, MD 20742
 sahanl@umd.edu ismailc@umd.edu ulukus@umd.edu

Abstract—We develop a novel source coding strategy for sampling and monitoring of a Wiener process. For the encoding process, we employ a four level “quantization” scheme, which employs monotone function thresholds as opposed to fixed constant thresholds. Leveraging the hitting times of the Wiener process with these thresholds, we devise a sampling and encoding strategy which does not incur any quantization errors. We give analytical expressions for the mean squared error (MSE) and find the optimal source code lengths to minimize the MSE under this monotone function threshold scheme, subject to a sampling rate constraint.

I. INTRODUCTION

Throughout the literature, Wiener processes or rather Brownian motion, has become a fundamental unit of study whose horizons expand from governing the statistical dynamics of particles to modeling the behavior of contemporary applications such as stock markets and computer networking [1], [2]. Thus, the remote estimation of a Wiener process has become a widely pursued research interest in recent past.

One of the first papers which explicitly studies the remote estimation of Wiener processes is [3], where the authors find the optimal sampling policy that should be employed when monitoring a Wiener process through a random delay channel, to minimize the mean squared error (MSE). In there, the authors show that a zero-wait (ZW) sampling policy is not optimal, and hence introduce a sample-and-wait policy which is proven to be optimal. Thereby, this study has spawned a multitude of its variants including its extensions to different processes such Ornstein-Uhlenbeck (OU) process, different channel statistics and multiple sources [4]–[6]. All these works are interesting in their own right, however, they all share a fundamental drawback: They all involve the transmission of the real-valued samples of the process which cannot be practically implemented. Thus, the need for an appropriate quantization scheme arises naturally.

The work in [7] takes the need for quantization into account and provides analytical expressions for the MSE over a finite horizon. They resort to a high resolution quantization scheme and highlight the relations between the age process [8], quantization errors and the MSE. Their transmission strategy involves quantizing the difference of the Wiener process between two successive sampling points, where they show that this strategy leads to unbounded quantization errors as the time horizon is expanded. To circumvent this problem, they resort to a multi-level quantization error correcting scheme which involves quantizing the accumulated sum of quantization errors in multiple stages. However, this leads to another problem:

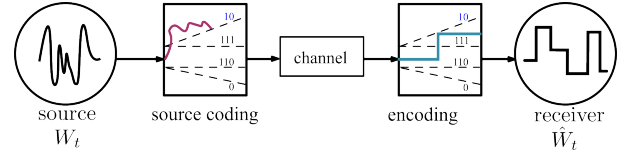


Fig. 1. System model.

In this procedure in order to transmit the N^i th sample, the transmitter needs to go through i quantization stages before transmission. Thus, the computational complexity of this scheme becomes unbounded in the infinite time horizon.

In our work, we aim to completely eliminate the quantization errors by employing an event-driven sampling strategy. In particular, we take samples only when our process hits a known threshold. Similar to [7], we too transmit the successive differences of the Wiener process between two sampling points. However, during the transmission of a sample, the Wiener process may very well be outside our predetermined thresholds. Hence, deviating from the norm, we employ two monotone function thresholds to effectively capture the process, if it goes outside the predetermined thresholds. This procedure is explained in detail in Section III. We exploit the hitting times of the process with these thresholds and provide analytical expressions for the MSE, $\mathbb{E}[\Delta_{mse}]$, of the process defined by,

$$\mathbb{E}[\Delta_{mse}] = \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\int_0^T (W_t - \hat{W}_t)^2 dt \right], \quad (1)$$

where W_t is the Wiener process and $\hat{W}_t$ is its estimate. Further, we develop a source coding scheme to minimize the MSE under a given sampling rate. The introduction of source coding scheme makes the transmission delay dependent on the signal, adding a new layer of complexity to the problem in [3], where the transmission delays were independent of the process.

To summarize our contributions: (i) We introduce a novel event-driven sampling strategy with monotone function thresholds for monitoring a Wiener process. (ii) We provide analytical expressions for the MSE under this monotone function thresholding scheme. (iii) We find the optimal source codes for the event-driven sampling and encoding strategy.

II. RELATED WORK

The work in [9] studies the problem of coding a Brownian motion in the interval of $(0, 1)$ so as to minimize the supremum of the quantization error under a given norm. They achieve

this by representing the Brownian motion through piecewise constant paths on intervals and encoding the exit times of the process from these intervals. However, since the exit times are also continuous, this too induces quantization errors. They overcome this by approximating the random variable associated with the exit times with an appropriate discrete random variable and treat its discrepancy as the quantization error. However, this study does not consider sampling or transmission processes into account.

The reference [10] looks into the problem of reconstructing a Wiener process from a quantized set of samples where they evaluate the trade-off between the sampling rate, bit rate and MSE. As opposed to per sample transmission, they sample the Wiener process multiple times and encode a vector of samples for transmission using a fixed number of bits each time.

In the realm of source coding, [11] analyzes the problem of sampling and transmission of a discrete random source so as to minimize the age of information. They show that Shannon codewords constructed with respect to a tilted version of the probability mass function (PMF) of the random source is close to optimal (only constant away). This idea is further expanded by the work in [12], which selectively encodes a subset of the most probable realizations instead of encoding all possible realizations of the random source.

The closest to our work is [13], where the authors employ a threshold based sampling scheme along with timing side information (TSI) and a buffer. However, as opposed to our system, they flood the channel with a sample whenever the process exceeds the threshold even during an ongoing transmission. Additionally, they encode and transmit the inter-sampling times and simply use a single bit to denote which threshold the process hits. An issue in this approach is the delays incurred due to the backlog of packets. As highlighted in [3], sampling when the channel is busy is not optimal when monitoring a Wiener process.

III. MONOTONE FUNCTION THRESHOLDS

Now, we describe our sampling and source coding scheme. Let W_t denote a Wiener process with unit variance. Assuming a bufferless system (no queue), a sampler will monitor this process and take samples when needed based on channel availability (i.e., sample only when channel is free). Once a sample is taken, it will be source coded and will be transmitted to the remote monitor (see Fig. 1). We assume that channel would incur a delay of one unit of time to transmit one bit. Suppose we take the n th sample at time S_n and it is coded into a length of L_n . Then, it would be delivered to the monitor at time $D_n = S_n + L_n$.

The monitor maintains an estimate of the Wiener process denoted by $\hat{W}_t$, which should be set to the value of the last received sample in order to minimize the MSE [3]. At each sampling instance, we will not be encoding and transmitting the actual value of the Wiener process, but rather its deviation from the previous sampling instance. In other words, we will be transmitting successive increments of W_t instead of the actual value of W_t . In particular, for the $(n + 1)$ th sample,

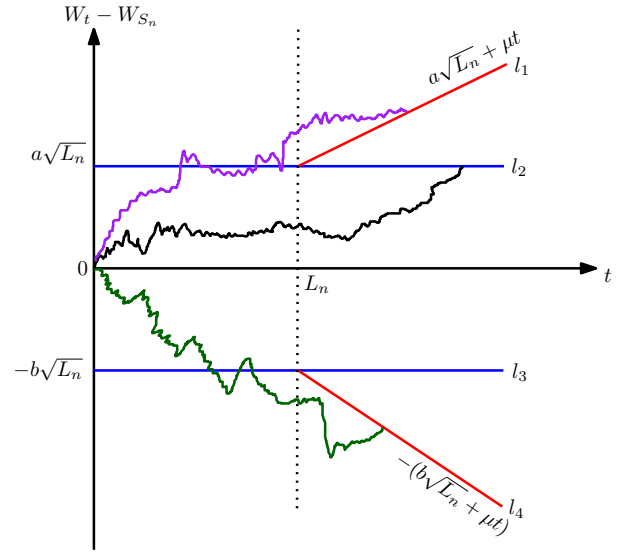


Fig. 2. Monotone function thresholding scheme.

we will be encoding and transmitting $Z_{n+1} = W_{S_{n+1}} - W_{S_n}$, instead of $W_{S_{n+1}}$. Hence, at time D_{n+1} , the monitor can update its estimator to $\hat{W}_{D_{n+1}} = W_{S_{n+1}} = \sum_{i=1}^{n+1} Z_i$ assuming, $W_0 = 0$. The monitor will maintain this update until the next sample is received.

To encode the successive increments, we employ the sampling and encoding strategy illustrated in Fig. 2. In particular, to take the $(n + 1)$ th sample, we will observe the process $W_t - W_{S_n}$ and wait until it hits one of the thresholds depicted in Fig. 2 after time $D_n = S_n + L_n$. In here a, b are some non-negative constants. Note that since the Wiener process has independent increments, we have taken $S_n = 0$ in Fig. 2. Let $X_n = W_{D_n} - W_{S_n}$. If $-b\sqrt{L_n} \leq X_n \leq a\sqrt{L_n}$, the next sample will be taken when $W_t - W_{S_n}$ hits either $a\sqrt{L_n}$ or $-b\sqrt{L_n}$. In this instance, we will be taking the next sample at the stopping time $\tilde{\tau}(X_n) = \min\{\tilde{\tau}_{a_n}, \tilde{\tau}_{b_n}\}$, where $\tilde{\tau}_{a_n} = \inf\{t : W_{t+D_n} - W_{S_n} = a\sqrt{L_n}\}$ and $\tilde{\tau}_{b_n} = \inf\{t : W_{t+D_n} - W_{S_n} = -b\sqrt{L_n}\}$. Moreover, $\tilde{\tau}(X_n)$ is a well-behaved stopping time (i.e., $\mathbb{E}[\tilde{\tau}(X_n)] < \infty$). If the process hits $a\sqrt{L_n}$, we will be sending a codeword of l_2 bits and if it hits $-b\sqrt{L_n}$ will be sending a codeword of l_3 bits. In either case, we will be transmitting and encoding which event occurred as opposed to the actual value of the sample. However, upon decoding the codeword, the monitor can unveil the actual value of the sample to be either $a\sqrt{L_n}$ or $-b\sqrt{L_n}$.

If $X_n > a\sqrt{L_n}$, we will wait until $W_t - W_{S_n}$ hits the increasing threshold defined by $a\sqrt{L_n} + \mu t$. In this case, we will be taking a sample at the stopping time which is defined by $\hat{\tau}_a(X_n) = \inf\{t : W_{t+D_n} - W_{S_n} = a\sqrt{L_n} + \mu t\}$. This stopping time is equivalent in distribution to a hitting time of a standard Brownian motion having a drift of $\mu > 0$, with a positive threshold. Therefore, this is also a well-behaved stopping time (see Corollary 1). This event will be encoded by a separate codeword which will be transmitted to the monitor. Upon decoding this codeword, the receiver has the following pieces of information: At time D_{n+1} , it can

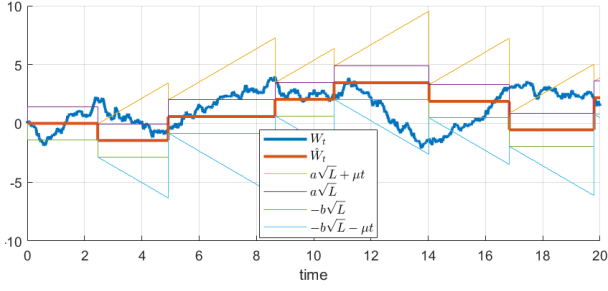


Fig. 3. A sample path with $a = b = \mu = 1$ and $L_n = L = 2$.

deduce the exact sampling time of the $(n+1)$ th sample as $S_{n+1} = D_{n+1} - L_{n+1}$ and thereby $\hat{\tau}_a(X_n)$ since it knows D_n . Next, from the received codeword, it knows that the sample was taken as result of process hitting the increasing threshold $a\sqrt{L_n} + \mu t$. Therefore, it can deduce the exact value of the successive increment using $Z_{n+1} = a\sqrt{L_n} + \mu\hat{\tau}_a(X_n)$ without inducing any quantization errors. In this case, we will be using l_1 bits to encode this event. The case $X_n < -b\sqrt{L_n}$ follows a similar treatment where we define the hitting time of the increment process with the decreasing threshold $-(b\sqrt{L_n} + \mu t)$ as $\hat{\tau}_b(X_n) = \inf\{t : W_{t+D_n} - W_{S_n} = -(b\sqrt{L_n} + \mu t)\}$ and we will use l_4 bits to encode this event.

In all four events, we will be transmitting the codeword associated with the event instead of encoding and transmitting the actual value of the Wiener process. Therefore, in general, we will be taking the $(n+1)$ th sample after D_n , at the stopping time defined $\tau(X_n) = \min\{\hat{\tau}(X_n), \hat{\tau}_a(X_n), \hat{\tau}_b(X_n)\}$ and encode and transmit the event that resulted in the stoppage. A sample path of how W_t and $\hat{W}_t$ change has been illustrated with Fig. 3. Further, we will see in the next section, that the probabilities of these four events are identical and independent of previous sampling instances (see Lemma 1) and hence the same codebook can be used for encoding all the samples.

Remark 1 When $\mu = 0$ (i.e., if we have only the constant thresholds instead of monotone function thresholds), if the process $X_n > a\sqrt{L_n}$ or $X_n < -b\sqrt{L_n}$, then the expected hitting time of the increment process with one of the constant thresholds will be infinite. Thus, the use of monotone function thresholds is necessary for a stable sampling procedure. Moreover, μt can be replaced with any increasing function $f(t)$, which leads to a well-behaved stopping time, e.g., $f(t) = \mu t^2$ or $f(t) = e^{\mu t}$, etc.

IV. MAIN RESULTS

In this section, we state the main analytical results when implementing the monotone function thresholding scheme. The very first result states that under the described scheme, since we appropriately scale the thresholds based on our previous source coding length, the resulting generation of source coding lengths will be i.i.d. The proof of Lemma 1 is given in Appendix A.

Lemma 1 The source coding lengths are generated independently and identically.

Next, using Lemma 1 and the optional stopping theory, we can find $\mathbb{E}[\Delta_{mse}]$ as stated in Theorem 1.

Theorem 1 The MSE of the Wiener process under the monotone function thresholding scheme is given by,

$$\mathbb{E}[\Delta_{mse}] = \frac{\mathbb{E}[Y^4] + 6\mathbb{E}[C(Y)Y^2]}{6\mathbb{E}[\tau(X_L) + L]}, \quad (2)$$

where $Y = X_L + B(\tau(X_L))$, $B(t)$ is a standard Brownian motion rooted at zero, $C(Y)$ is the source coding associated with Y , $X_L \sim \mathcal{N}(0, L)$ and L is the random variable associated with code lengths.

The proof of Theorem 1 can be found in Appendix B. Next, let p_i be the probability that source coding length l_i is generated for the next sample (see Appendix A for a formal description of p_i s). Then, for a given set of thresholds and for large μ , Theorem 2 gives the closed-form expression for $\mathbb{E}[\Delta_{mse}]$. For the rest of the paper, we will assume μ is sufficiently large.

Theorem 2 For sufficiently large μ , $\mathbb{E}[\Delta_{mse}]$ is given by,

$$\mathbb{E}[\Delta_{mse}] = K \frac{\mathbb{E}_P[L^2]}{\mathbb{E}_P[L]} + \mathbb{E}_{\tilde{P}}[L], \quad (3)$$

where

$$K = \frac{3 + p_2 a^4 + p_3 b^4 - \tilde{X}}{6(p_2 a^2 + p_3 b^2 + \tilde{A} + \tilde{B})}, \quad (4)$$

$$\tilde{A} = \frac{1}{\sqrt{2\pi}} \int_a^\infty x^2 e^{-\frac{x^2}{2}} dx, \quad (5)$$

$$\tilde{B} = \frac{1}{\sqrt{2\pi}} \int_b^\infty x^2 e^{-\frac{x^2}{2}} dx, \quad (6)$$

$$\tilde{X} = \frac{1}{\sqrt{2\pi}} \int_{-b}^a x^4 e^{-\frac{x^2}{2}} dx, \quad (7)$$

and $P = \{p_1, p_2, p_3, p_4\}$ and $\tilde{P} = \{\frac{\tilde{A}}{D}, \frac{p_2 a^2}{D}, \frac{p_3 b^2}{D}, \frac{\tilde{B}}{D}\}$ with $D = p_2 a^2 + p_3 b^2 + \tilde{A} + \tilde{B}$, are PMFs on source coding lengths $\{l_1, l_2, l_3, l_4\}$. Moreover, the sampling rate (SR) of the scheme is given by,

$$SR = \frac{1}{\mathbb{E}[\tau(X_L) + L]} = D\mathbb{E}_P[L]. \quad (8)$$

Remark 2 Note that the expression given in (3) is analogous to the average of age of a process whose inter arrival times between updates are governed by the random variable L similar to the work in [11].

Remark 3 When μ is sufficiently large and $l_i = 2$ for all i , our scheme effectively reduces to the optimal sampling scheme of Wiener process given in [3] under a constant transmission delay of 2.

Remark 4 Note that μ can be as large as desired, however it cannot be infinite, since this will render the process of estimating the actual value of the Wiener process using the

time difference of the samples useless. No information about the actual value of the increments can be transmitted if we set $\mu = \infty$.

Remark 5 From (3), it is evident that $\mathbb{E}[\Delta_{mse}]$ scales linearly with the bit transmission delay.

Remark 6 If the Wiener process has a variance of σ^2 , replacing a , b and μ with σa , σb and $\sigma \mu$, respectively, linearly scales $\mathbb{E}[\Delta_{mse}]$ with respect to σ^2 .

V. OPTIMAL SOURCE CODES

In this section, we will find the optimal source coding lengths to minimize $\mathbb{E}[\Delta_{mse}]$ subject to a sampling rate constraint. Now, the task at hand is two folds. First, for given values of a and b , we need to find a set of prefix codeword lengths that minimizes (3). This is equivalent to minimizing (3) with respect to l_i subject to the Kraft's inequality. Next, we need to find the optimal values for the thresholds. Thus, our problem can be posed as the following optimization problem,

$$\begin{aligned} \min_{a, b > 0, l_i \in \mathbb{N}} \quad & \frac{K\mathbb{E}_P[L^2]}{\mathbb{E}_P[L]} + \mathbb{E}_{\tilde{P}}[L] \\ \text{s.t.} \quad & \sum_{i=1}^4 2^{-l_i} \leq 1, \\ & \mathbb{E}_P[L] \geq \frac{1}{Df_{max}}, \end{aligned} \quad (9)$$

where f_{max} is the maximum sampling rate allowed.

Unlike in traditional source coding problems where the objective is a linear function of coding lengths, in here, the objective function is quadratic with respect to source coding lengths. Now, instead of solving the optimization problem (9) exactly, we will solve for a relaxed version of the optimization problem by considering $l_i \in \mathbb{R}$ instead of restricting to non-negative integers. When assuming $l_i \in \mathbb{R}$, there is no real merit in considering two distinct values for a and b due to the symmetry of the problem. Therefore, we will consider $a = b$ for the rest of problem. Further, $a = b$ implies that $l_1 = l_4$, $l_2 = l_3$, $p_1 = p_4$ and $p_2 = p_3$. For a fixed a , we will first find the best source coding lengths and later optimize with respect to a . Thus, the problem at hand, reduces to the following,

$$\begin{aligned} \min_{l_1, l_2 \in \mathbb{R}} \quad & \frac{K\mathbb{E}_P[L^2]}{\mathbb{E}_P[L]} + \mathbb{E}_{\tilde{P}}[L] \\ \text{s.t.} \quad & \sum_{i=1}^2 2^{-l_i} \leq \frac{1}{2}, \\ & \mathbb{E}_P[L] \geq \frac{1}{Df_{max}}. \end{aligned} \quad (10)$$

Then, using the Dinkelbach method [14], we will linearize the fractional objective function as follows,

$$\min_{l_i \in \mathbb{R}} \quad K\mathbb{E}_P[L^2] + \mathbb{E}_P[L]\mathbb{E}_{\tilde{P}}[L] - \theta\mathbb{E}_P[L] \quad (11)$$

where $\theta \in \mathbb{R}$. This is optimized with respect to the same constraints as in (10). Let $J(\theta)$ denote the optimal value of

the above optimization and J denote the optimal value of (10). Then, it is well-known that when $J(\theta) \leq 0 \iff \theta \leq J$ [3]. The above optimization problem satisfies the linear independence constraint qualification (LIQC) and hence the Karush-Khun-Tucker (KKT) conditions give the necessary conditions for optimality. Let the Lagrangian of the above optimization problem be defined as,

$$\begin{aligned} \mathcal{L}(l, \lambda) = & K\mathbb{E}_P[L^2] + \mathbb{E}_P[L]\mathbb{E}_{\tilde{P}}[L] - \theta\mathbb{E}_P[L] \\ & + \gamma \left(\frac{1}{Df_{max}} - \mathbb{E}_P[L] \right) + \lambda \left(\sum_{i=1}^2 2^{-l_i} - \frac{1}{2} \right), \end{aligned} \quad (12)$$

where $\gamma, \lambda \geq 0$ are the Lagrangian multipliers. Then, the KKT conditions are given by,

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial l_i} = & 4Kl_i p_i + 2\tilde{p}_i \mathbb{E}_P[L] + 2p_i \mathbb{E}_{\tilde{P}}[L] - 2(\theta + \gamma)p_i \\ & - \ln 2 \lambda 2^{-l_i} = 0, \end{aligned} \quad (13)$$

with the complementary slackness conditions given by,

$$\lambda \left(\sum_{i=1}^2 2^{-l_i} - \frac{1}{2} \right) = 0, \quad (14)$$

$$\gamma \left(\frac{1}{Df_{max}} - \mathbb{E}_P[L] \right) = 0. \quad (15)$$

In here, $\tilde{p}_i$ is simply the i th element of the PMF $\tilde{P}$ defined in Theorem 2. Next, Theorem 3 states an important characteristic of the optimal source coding lengths.

Theorem 3 For a given threshold, the optimal source coding lengths must satisfy either the Kraft inequality or the sampling rate constraint with equality.

Remark 7 As a consequence of Theorem 3, if $f_{max} = \infty$, then the optimal source coding lengths must satisfy the Kraft inequality with equality.

The proof of Theorem 3 is given in Appendix D. Now, to solve the $J(\theta)$ optimization, the objective function of (11) can be expressed as a quadratic program as follows,

$$\min_{l \in \mathbb{R}^2} \quad l^T Q l - q_\theta^T l \quad (16)$$

where $l = [l_1, l_2]^T$, $q_\theta = [2\theta p_1, 2\theta p_2]^T$ and Q is defined as,

$$Q = 2 \begin{bmatrix} Kp_1 + 2p_1\tilde{p}_1 & p_1\tilde{p}_2 + p_2\tilde{p}_1 \\ p_2\tilde{p}_1 + p_1\tilde{p}_2 & K + 2\tilde{p}_2 p_2 \end{bmatrix}. \quad (17)$$

The quadratic program must be solved subject to the same constraints in (10). Note that Q is independent of θ and only dependent on a . A simple numerical verification shows that Q is positive semi-definite for all $a > 0$. Since the constraints in the optimization (10) are also convex, the above quadratic program can be easily solved. This gives us $J(\theta)$. Now, the optimal θ which is the solution for $J(\theta) = 0$, can be found using a simple bisection search. Then, finally, we perform a one-dimensional exhaustive search to find the optimal a .

VI. NUMERICAL RESULTS

In this section, we validate our analysis via simulations and compare our proposed scheme with several benchmark schemes. For simulations, we discretize time t into time intervals of length $\varepsilon = 10^{-2}$, and we construct the Wiener process at these discretized time indices as $W_{t_i} = W_{t_{i-1}} + r$, where $r \sim \mathcal{N}(0, \varepsilon)$. In each time index, if there is no ongoing transmission, we check if the Wiener process has reached one of the thresholds described in Section III. In the event that it crosses one of the given thresholds, the event is source coded with L bits and transmitted over $\text{round}(L/\varepsilon)$ time indices. In here, the round operation ensures that the transmission duration takes an integer multiple of our discrete time interval ε . After transmission, the transmitted event is decoded and the estimate $\hat{W}_t$ is updated. This process is repeated over the time horizon of $T = 10^5$ (i.e., $\text{round}(T/\varepsilon)$ time indices).

Fig. 4 illustrates the comparison of our proposed scheme with the benchmark schemes when there is no sampling rate constraint. In line with our symmetry argument in Section V, we fix $a = b$, $\ell_1 = \ell_4$, and $\ell_2 = \ell_3$. Then, we obtain the optimal ℓ_1 and ℓ_2 that minimize the MSE for each threshold a . The *uniform source code* is a special case of our monotone thresholding scheme, where each event is encoded with 2 bits (i.e., $\ell_i = 2, \forall i$). Therefore, in this case, each transmission takes 2 time units. In the *ideal sampling* scheme we assume that the source is able to transmit the real-value without any source coding in unit time. In here, the source samples and transmits whenever the Wiener process exceeds the threshold a when the channel is free. As suggested in [3], even in the absence of a sampling rate constraint, zero-wait sampling (i.e., $a = 0$) is not optimal for the described benchmark schemes. Therefore, as illustrated, the minimum MSE is achieved with a non-zero a for these benchmark schemes. On the other hand, in the absence of a sampling rate constraint, $a = 0$ is optimum for the optimal source coding scheme.

Fig. 4(a), compares MSE values of the three schemes where the solid curves represent the analytical results, and dots correspond to simulation results. First, we observe that the analytical curves fit with the simulation results, verifying our analysis. Additionally, we observe that the MSE performance of the optimal source coding scheme is bounded by the other two schemes. In general, when a is small, p_1 is also small. Therefore, the MSE is minimized by selecting ℓ_1 close to 1, and setting ℓ_2 to a large value. Notice that when threshold a approaches 0, it makes $p_1 = 0$ and (3) reduces to $\frac{3}{2}\ell_1$. This is minimized by setting $\ell_1 = 1$ and $\ell_2 = \infty$ which also sets the sampling rate to 1. At this point the ideal sampling scheme and our optimal source coding scheme are identical and hence the MSEs are equivalent.

As Fig. 4 suggests, small a values cause a high sampling rate for the optimal source coding in the absence of a sampling rate constraint. Therefore, in Fig 5, we illustrate the optimum MSE values under different sampling rate constraints and their corresponding sampling rates in Fig. 5(a) and Fig. 5(b), respectively. As highlighted in Theorem 3, optimal source

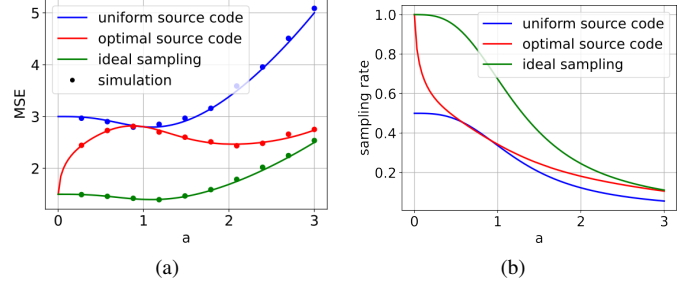


Fig. 4. Comparison of optimum source code scheme under no sample rate constraint with benchmark schemes. Curves are obtained from analytical expressions, and dots represent simulation results.

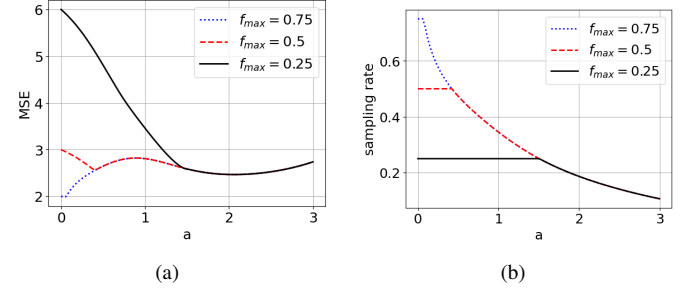


Fig. 5. a) MSE and b) sampling rate under different sampling rate constraints.

coding scheme must satisfy at least one of the sampling rate and Kraft's constraints with equality. Therefore, in Fig 5(b), we can identify two regions for a , one in which the sampling rate constraint is active (for small a) and the other in which the Kraft inequality is active (for large a). At the intersecting point of these regions, both of these constraints are active. In addition, when a is small, the MSE and the sampling rate values are mainly determined by ℓ_1 . Since the sampling rate constraint is active in this region, ℓ_1 cannot be chosen arbitrarily close to 1 anymore, and hence the optimal a value is non-zero now.

VII. CONCLUSION

We introduced a novel sampling and encoding strategy for monitoring a Wiener process. For this scheme, we have found the optimal source code lengths which suggests that a zero-wait sampling policy is optimal in the absence of a sampling rate constraint. In the presence of a sampling rate constraint, a non-zero threshold scheme with monotone function thresholds is deemed to be optimal. The effect of multiple thresholds along with monotone function thresholds is an interesting research direction which is left for future work.

APPENDIX A PROOF OF LEMMA 1

Let L_n be the source coding length of the previous sample. Let $X_n = W_{D_n} - W_{S_n}$. If $X_n > a\sqrt{L_n}$, then $L_{n+1} = \ell_1$ and if $X_n < -b\sqrt{L_n}$ then $L_{n+1} = \ell_4$. If $X_n = x$ for $x \in (-b\sqrt{L_n}, a\sqrt{L_n})$, then the $W_t - S_n$ will hit the upper constant threshold with probability, $\frac{(x+b\sqrt{L_n})}{\sqrt{L_n}(a+b)}$ and would hit probability

$\frac{(a\sqrt{L_n})-x}{\sqrt{L_n}(a+b)}$. If it hits the upper threshold we will have $L_{n+1} = l_2$ and if hits the lower threshold we will have $L_{n+1} = l_3$. Also note that $X_n \sim \mathcal{N}(0, L_n)$. Therefore, L_{n+1} will take the values in $\{l_1, l_2, l_3, l_4\}$ with the following probabilities,

$$\begin{aligned}\mathbb{P}(L_{n+1} = l_1 | L_n) &= \frac{1}{\sqrt{2\pi L_n}} \int_{a\sqrt{L_n}}^{\infty} e^{-\frac{x^2}{2L_n}} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_a^{\infty} e^{-\frac{x^2}{2}} dx \\ &= p_1,\end{aligned}\quad (18)$$

$$\begin{aligned}\mathbb{P}(L_{n+1} = l_2 | L_n) &= \frac{1}{\sqrt{2\pi L_n}} \int_{-b\sqrt{L_n}}^{a\sqrt{L_n}} \frac{(x + b\sqrt{L_n})}{\sqrt{L_n}(a+b)} e^{-\frac{x^2}{2L_n}} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-b}^a \frac{(x+b)}{(a+b)} e^{-\frac{x^2}{2}} dx \\ &= p_2,\end{aligned}\quad (19)$$

$$\begin{aligned}\mathbb{P}(L_{n+1} = l_3 | L_n) &= \frac{1}{\sqrt{2\pi L_n}} \int_{-b\sqrt{L_n}}^{a\sqrt{L_n}} \frac{(a\sqrt{L_n} - x)}{\sqrt{L_n}(a+b)} e^{-\frac{x^2}{2L_n}} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-b}^a \frac{(a-x)}{(a+b)} e^{-\frac{x^2}{2}} dx \\ &= p_3,\end{aligned}\quad (20)$$

$$\begin{aligned}\mathbb{P}(L_{n+1} = l_4 | L_n) &= \frac{1}{\sqrt{2\pi L_n}} \int_{-\infty}^{-b\sqrt{L_n}} e^{-\frac{x^2}{2L_n}} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-b} e^{-\frac{x^2}{2}} dx \\ &= p_4.\end{aligned}\quad (21)$$

As seen above, the conditional probabilities of L_{n+1} are independent of L_n , thus proving the result.

APPENDIX B PROOF OF THEOREM 1

Let us first state an important lemma derived from the optional stopping theory of Wiener process whose proof is available in [3] and [15].

Lemma 2 *Let τ be a stopping time of the Wiener process with $\mathbb{E}[\tau] < \infty$. If either $\sup_{0 \leq t \leq \tau} |W_t| < \infty$ or if $\mathbb{E}[\tau^2] < \infty$, then we have from optional stopping theorem and Doob's L^p maximal inequality,*

$$\mathbb{E} \left[\int_0^\tau W_t^2 dt \right] = \frac{\mathbb{E}[W_\tau^4] - \mathbb{E}[W_0^4]}{6}. \quad (22)$$

Let $X_n = W_{D_n} - W_{S_n}$. Then, the MSE from D_n to D_{n+1} can be obtained as follows,

$$\begin{aligned}&\mathbb{E} \left[\int_{D_n}^{D_{n+1}} (W_t - \hat{W}_t)^2 dt \right] \\ &= \mathbb{E} \left[\int_{D_n}^{D_n + \tau(X_n) + L_{n+1}} (W_t - W_{D_n} + W_{D_n} - W_{S_n})^2 dt \right]\end{aligned}\quad (23)$$

$$= \mathbb{E} \left[\int_0^{\tau(X_n) + L_{n+1}} (W_{t+D_n} - W_{D_n} + W_{D_n} - W_{S_n})^2 dt \right] \quad (24)$$

$$= \mathbb{E} \left[\int_0^{\tau(X_n) + L_{n+1}} (B(t) + X_n)^2 dt \right]. \quad (25)$$

Since $X_n \sim \mathcal{N}(0, L_n)$ and L_n are i.i.d., from renewal reward theorem (RRT), we have the following,

$$\mathbb{E}[\Delta_{mse}] = \frac{\mathbb{E} \left[\int_0^{\tau(X) + \tilde{L}} (B(t) + X)^2 dt \right]}{\mathbb{E}[\tau(X) + L]}, \quad (26)$$

where $X \sim \mathcal{N}(0, L)$, L is the discrete random variable associated with the source code lengths of the previous sample and $\tilde{L}$ is the source coding length of the next sample which has the same distribution as L .

Now, let $\tau(X) = \tau_x$ and $Y_t = B(t) + X$. Then, from Lemma 2, we have,

$$\begin{aligned}&6\mathbb{E} \left[\int_0^{\tau(X) + \tilde{L}} (B(t) + X)^2 dt \right] \\ &= 6\mathbb{E} \left[\int_0^{\tau_x + \tilde{L}} Y_t^2 dt \right]\end{aligned}\quad (27)$$

$$= \mathbb{E}[Y_{\tau_x + \tilde{L}}^4] - \mathbb{E}[X^4] \quad (28)$$

$$= \mathbb{E}[(Y_{\tau_x + \tilde{L}} - Y_{\tau_x} + Y_{\tau_x})^4] - 3\mathbb{E}[L^2] \quad (29)$$

$$\begin{aligned}&= \mathbb{E}[(Y_{\tau_x + \tilde{L}} - Y_{\tau_x})^4] + 6\mathbb{E}[(Y_{\tau_x + \tilde{L}} - Y_{\tau_x})^2 Y_{\tau_x}^2] \\ &\quad + \mathbb{E}[Y_{\tau_x}^4] - 3\mathbb{E}[L^2]\end{aligned}\quad (30)$$

$$= 3\mathbb{E}[\tilde{L}^2] + 6\mathbb{E}[\tilde{L}Y_{\tau_x}^2] + \mathbb{E}[Y_{\tau_x}^4] - 3\mathbb{E}[L^2] \quad (31)$$

$$= \mathbb{E}[Y_{\tau_x}^4] + 6\mathbb{E}[C(Y_{\tau_x})Y_{\tau_x}^2]. \quad (32)$$

APPENDIX C PROOF OF THEOREM 2

Before proceeding with the computation of $\mathbb{E}[\Delta_{mse}]$, we will first state two useful properties of the considered stopping times, which will be helpful towards proving the main result. The proof of Lemma 3 can be found in [15].

Lemma 3 *For $\mu > 0$, let τ_c be the hitting time of a Brownian process $\mu t + B(t)$ rooted at zero, with level $c > 0$. Then the Laplace transform $\Psi(\lambda)$ of the distribution of τ_c and the moments of τ_c will be given by,*

$$\Psi(\lambda) = \mathbb{E}[e^{-\lambda\tau_c}] = e^{-c(\sqrt{\mu^2 + 2\lambda} - \mu)}, \quad (33)$$

$$\mathbb{E}[\tau_c^k] = (-1)^k \frac{\partial^k \Psi}{\partial \lambda^k} \Big|_{\lambda=0}. \quad (34)$$

Corollary 1 *For $\mu > 0$, let τ_c be the hitting time of a Brownian process $\mu t + B(t)$ rooted at zero, with level $c > 0$. Then, the first 4 moments will be given by,*

$$\mathbb{E}[\tau_c] = \frac{c}{\mu}, \quad (35)$$

$$\mathbb{E}[\tau_c^2] = \frac{c^2}{\mu^2} + \frac{c}{\mu^3}, \quad (36)$$

$$\mathbb{E}[\tau_c^3] = \frac{c^3}{\mu^3} + \frac{3c^2}{\mu^4} + \frac{3c}{\mu^5}, \quad (37)$$

$$\mathbb{E}[\tau_c^4] = \frac{c^4}{\mu^4} + \frac{6c^3}{\mu^5} + \frac{15c^2}{\mu^6} + \frac{15c}{\mu^7}. \quad (38)$$

Next, define by A_k and B_k the following integrals,

$$A_k = \frac{1}{\sqrt{2\pi}} \int_a^\infty (x-a)^k e^{-\frac{x^2}{2}} dx, \quad (39)$$

$$B_k = \frac{1}{\sqrt{2\pi}} \int_b^\infty (x-b)^k e^{-\frac{x^2}{2}} dx. \quad (40)$$

Now, the MSE can be obtained as follows,

$$\begin{aligned} \mathbb{E}[C(Y)Y^2|X > a\sqrt{L}, X, L] \\ = l_1 \mathbb{E}[(a\sqrt{L} + \mu\tau_x)^2] \end{aligned} \quad (41)$$

$$= l_1(a^2L + 2a\mu\sqrt{L}\mathbb{E}[\tau_x] + \mu^2\mathbb{E}[\tau_x^2]), \quad (42)$$

where $\tau(X) = \tau_x$ is used for brevity. Now, note that τ_x for $x > a\sqrt{L}$ is equal in distribution to the hitting time of a Brownian motion with a drift of μ with level $x - a\sqrt{L}$. Therefore, we further have that,

$$\begin{aligned} \mathbb{E}[C(Y)Y^2|X > a\sqrt{L}, X, L] \\ = l_1 \left(a^2L + (2a\sqrt{L})(x_L - a\sqrt{L}) \right. \\ \left. + (x_L - a\sqrt{L})^2 + \frac{(x_L - a\sqrt{L})}{\mu} \right), \end{aligned} \quad (43)$$

where x_L is a realization of X given L . Now, accounting for the distribution of X yields,

$$\begin{aligned} \mathbb{E}[\mathbb{1}\{X > a\sqrt{L}\}C(Y)Y^2|L] \\ = l_1 \left(p_1a^2L + (2a\sqrt{L})\sqrt{L}A_1 + LA_2 + \frac{\sqrt{L}A_1}{\mu} \right) \end{aligned} \quad (44)$$

$$= \underbrace{(p_1a^2 + 2aA_1 + A_2)}_{\tilde{A}} l_1L + l_1 \frac{A_1\sqrt{L}}{\mu} \quad (45)$$

$$= l_1\tilde{A}L + l_1 \frac{A_1\sqrt{L}}{\mu}. \quad (46)$$

Similarly, for $X < -b\sqrt{L}$, we have,

$$\begin{aligned} \mathbb{E}[\mathbb{1}\{X < -b\sqrt{L}\}C(Y)Y^2|L] \\ = \underbrace{(p_1b^2 + 2bB_1 + B_2)}_{\tilde{B}} l_4L + l_4 \frac{B_1\sqrt{L}}{\mu} \end{aligned} \quad (47)$$

$$= l_4\tilde{B}L + l_4 \frac{B_1\sqrt{L}}{\mu}. \quad (48)$$

Now, considering the case $X \in (-b\sqrt{L}, a\sqrt{L})$, we have,

$$\begin{aligned} \mathbb{E}[\mathbb{1}\{-b\sqrt{L} \leq X \leq a\sqrt{L}\}C(Y)Y^2|L] \\ = L(p_2l_2a^2 + p_3l_3b^2). \end{aligned} \quad (49)$$

Therefore, we have,

$$\mathbb{E}[C(Y)Y^2] = \mathbb{E}\left[l_1\tilde{A}L + l_1 \frac{A_1\sqrt{L}}{\mu} + l_4\tilde{B}L + l_4 \frac{B_1\sqrt{L}}{\mu}\right]$$

$$+ (p_2l_2a^2 + p_3l_3b^2)L \quad (50)$$

$$\begin{aligned} &= (l_1\tilde{A} + l_4\tilde{B} + p_2l_2a^2 + p_3l_3b^2)\mathbb{E}_P[L] \\ &\quad + \frac{(l_1A_1 + l_4B_1)}{\mu}\mathbb{E}_P[\sqrt{L}]. \end{aligned} \quad (51)$$

Similarly, $\mathbb{E}[Y^4]$ can also be evaluated as follows,

$$\begin{aligned} \mathbb{E}[Y^4|X > a\sqrt{L}, X, L] \\ = (a\sqrt{L} + \mu\tau_x)^4 \end{aligned} \quad (52)$$

$$\begin{aligned} &= a^4L^2 + 4a^3L\sqrt{L}\mu\mathbb{E}[\tau_x] + 6a^2L\mu^2\mathbb{E}[\tau_x^2] \\ &\quad + 4a\sqrt{L}\mu^3\mathbb{E}[\tau_x^3] + \mu^4\mathbb{E}[\tau_x^4]. \end{aligned} \quad (53)$$

Now, using Corollary 1 and accounting for the distribution of X yields,

$$\begin{aligned} \mathbb{E}[\mathbb{1}\{X > a\sqrt{L}\}Y^4|L] \\ = p_1a^4L^2 + 4a^3L\sqrt{L}(A_1\sqrt{L}) + 6a^2L \left(A_2L + \frac{A_1\sqrt{L}}{\mu} \right) \\ + 4a\sqrt{L} \left(A_3L\sqrt{L} + \frac{3A_2L}{\mu} + \frac{3A_1\sqrt{L}}{\mu^2} \right) \\ + A_4L^2 + \frac{6A_3L\sqrt{L}}{\mu} + \frac{15A_2L}{\mu^2} + \frac{15A_1\sqrt{L}}{\mu^3}. \end{aligned} \quad (54)$$

Similarly, we have,

$$\begin{aligned} \mathbb{E}[\mathbb{1}\{X < -b\sqrt{L}\}Y^4|L] \\ = p_4b^4L^2 + 4b^3L\sqrt{L}(B_1\sqrt{L}) + 6b^2L \left(B_2L + \frac{B_1\sqrt{L}}{\mu} \right) \\ + 4b\sqrt{L} \left(B_3L\sqrt{L} + \frac{3B_2L}{\mu} + \frac{3B_1\sqrt{L}}{\mu^2} \right) \\ + B_4L^2 + \frac{6B_3L\sqrt{L}}{\mu} + \frac{15B_2L}{\mu^2} + \frac{15B_1\sqrt{L}}{\mu^3}. \end{aligned} \quad (55)$$

Also, we have,

$$\mathbb{E}[\mathbb{1}\{-b\sqrt{L} \leq X \leq a\sqrt{L}\}Y^4|L] = L^2(p_2a^4 + p_3b^4). \quad (56)$$

Hence, $\mathbb{E}[Y^4]$ can be computed by summing the above three equations and taking the expectation with respect to L .

Now, to evaluate the denominator of the (26), we have from Corollary 1,

$$\mathbb{E}[\mathbb{1}\{X > a\sqrt{L}\}\tau(X)] = \frac{\sqrt{L}A_1}{\mu}, \quad (57)$$

$$\mathbb{E}[\mathbb{1}\{X < -b\sqrt{L}\}\tau(X)] = \frac{\sqrt{L}B_1}{\mu}, \quad (58)$$

and from Wald's identity for the Wiener process, we have,

$$\begin{aligned} \mathbb{E}\left[\mathbb{1}\{-b\sqrt{L} \geq X \leq a\sqrt{L}\}\tau(X)\right] \\ = \mathbb{E}\left[\mathbb{E}[\tau(X)|X, \mathbb{1}\{-b\sqrt{L} \geq X \leq a\sqrt{L}\}]\right] \end{aligned} \quad (59)$$

$$= \mathbb{E}\left[\mathbb{E}[Y^2 - X^2|X, \mathbb{1}\{-b\sqrt{L} \geq X \leq a\sqrt{L}\}]\right] \quad (60)$$

$$= (p_2a^2 + p_3b^2)L - \frac{L}{\sqrt{2\pi}} \int_{-b}^a x^2 e^{-\frac{x^2}{2}} dx. \quad (61)$$

(62) Now, we can obtain $J(\theta)$ as follows,

$$\begin{aligned} J(\theta) &= \sum_{i=1}^4 p_i \left(1 + \frac{p_i - \tilde{p}_i}{2Kp_i}\right)^2 (\mathbb{E}_P[L])^2 + \mathbb{E}_P[L](\mathbb{E}_{\tilde{P}}[L] - \theta) \\ &= \sum_{i=1}^4 p_i \left(1 + \frac{p_i - \tilde{p}_i}{2Kp_i}\right)^2 (\mathbb{E}_P[L])^2 - (2K+1)(\mathbb{E}_P[L])^2 \\ &= \underbrace{\left(\sum_{i=1}^4 p_i \left(1 + \frac{p_i - \tilde{p}_i}{2Kp_i}\right)^2 - (2K+1)\right)}_{\tilde{K}} (\mathbb{E}_P[L])^2. \end{aligned} \quad (76)$$

A simple numerical verification shows that $\tilde{K} < 0$, for all $a \geq 0$. Since at the optimal θ , we have $J(\theta) = 0$, this leads to a contradiction as $\mathbb{E}_P[L] > 0$.

REFERENCES

- [1] M. F. M. Osborne. Brownian motion in the stock market. *Operations Research*, 7(2):145–173, April 1959.
- [2] L. M. Wein. Brownian networks with discretionary routing. *Operations Research*, 39(2):322–340, March 1991.
- [3] Y. Sun, Y. Polyanskiy, and E. Uysal. Sampling of the Wiener process for remote estimation over a channel with random delay. *IEEE Transactions on Information Theory*, 66(2):1118–1135, February 2020.
- [4] S. Liyanarachchi and S. Ulukus. The role of early sampling in age of information minimization in the presence of ACK delays. *IEEE Transactions on Information Theory*, 70(9):6665–6678, September 2024.
- [5] H. Tang, Y. Sun, and L. Tassiulas. Sampling of the Wiener process for remote estimation over a channel with unknown delay statistics. *IEEE/ACM Transactions on Networking*, 32(3):1920–1935, June 2024.
- [6] K. Banawan, A. Arafa, and K. G. Seddik. Timely multi-process estimation over erasure channels with and without feedback: Signal-independent policies. *IEEE Journal on Selected Areas in Information Theory*, 4:607–623, November 2023.
- [7] S. Hu, H. Hui, and W. Chen. Monitoring the Brownian motion in real time: A rate-distortion perspective. In *IEEE SPAWC*, September 2021.
- [8] R. D. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus. Age of information: An introduction and survey. *IEEE Journal on Selected Areas in Communication*, 39(5):1183–1210, May 2021.
- [9] F. Poloczek and F. Ciucu. Coding of Brownian motion by quantization of exit times. In *Allerton Conference*, October 2012.
- [10] A. Kipnis, A. J. Goldsmith, and Y. C. Eldar. The distortion-rate function of sampled Wiener processes. *IEEE Transactions on Information Theory*, 65(1):482–499, January 2019.
- [11] P. Mayekar, P. Parag, and H. Tyagi. Optimal source codes for timely updates. *IEEE Transactions on Information Theory*, 66(6):3714–3731, June 2020.
- [12] M. Bastopcu, B. Buyukates, and S. Ulukus. Selective encoding policies for maximizing information freshness. *IEEE Transactions on Communications*, 69(9):5714–5726, September 2021.
- [13] S. Yu, W. Chen, and H. V. Poor. TSI-aided real-time monitoring of Brownian motions: A rate-latency-distortion perspective. In *IEEE Globecom*, December 2022.
- [14] W. Dinkelbach. On nonlinear fractional programming. *IEEE Transactions on Communications*, 13(7):492–498, March 1967.
- [15] R. Serfozo. *Basics of Applied Stochastic Process*. Springer, 2009.

Therefore, we have,

$$\begin{aligned} \mathbb{E}[\tau(X) + L] &= (p_2a^2 + p_3b^2 + (\tilde{A} + \tilde{B}))\mathbb{E}_P[L] \\ &\quad + \frac{(A_1 + B_1)\mathbb{E}_P[\sqrt{L}]}{\mu}. \end{aligned} \quad (63)$$

Now if set μ to a large value, we have,

$$\begin{aligned} \mathbb{E}[Y^4] &= \left((p_1a^4 + 4a^3A_1 + 6a^2A_2 + 4aA_3 + A_4) \right. \\ &\quad \left. + (p_4b^4 + 4b^3B_1 + 6b^2B_2 + 4bB_3 + B_4) \right. \\ &\quad \left. + (p_2a^4 + p_3b^4) \right) \mathbb{E}_P[L^2] \end{aligned} \quad (64)$$

$$\begin{aligned} &= \left(\frac{1}{\sqrt{2\pi}} \int_a^\infty x^4 e^{-\frac{x^2}{2}} dx + \frac{1}{\sqrt{2\pi}} \int_b^\infty x^4 e^{-\frac{x^2}{2}} dx \right. \\ &\quad \left. + (p_2a^4 + p_3b^4) \right) \mathbb{E}_P[L^2] \end{aligned} \quad (65)$$

$$= \left(3 + p_2a^4 + p_3b^4 - \frac{1}{\sqrt{2\pi}} \int_{-b}^a x^4 e^{-\frac{x^2}{2}} dx \right) \mathbb{E}_P[L^2] \quad (66)$$

$$= (3 + p_2a^4 + p_3b^4 - \tilde{X})\mathbb{E}_P[L^2], \quad (67)$$

and

$$\mathbb{E}[C(Y)Y^2] = (l_1\tilde{A} + l_4\tilde{B} + p_2l_2a^2 + p_3l_3b^2)\mathbb{E}_P[L]. \quad (68)$$

Further, $\mathbb{E}[\tau(X) + L]$ simplifies to the following,

$$\mathbb{E}[\tau(X) + L] = (p_2a^2 + p_1b^2 + \tilde{A} + \tilde{B})\mathbb{E}_P[L]. \quad (69)$$

Thus, $\mathbb{E}[\Delta_{mse}]$ will be given by,

$$\begin{aligned} \mathbb{E}[\Delta_{mse}] &= \frac{(3 + p_2a^4 + p_3b^4 - \tilde{X})\mathbb{E}_P[L^2]}{6(p_2a^2 + p_3b^2 + \tilde{A} + \tilde{B})\mathbb{E}_P[L]} \\ &\quad + \frac{l_1\tilde{A} + l_4\tilde{B} + p_2l_2a^2 + p_3l_3b^2}{p_2a^2 + p_3b^2 + \tilde{A} + \tilde{B}} \end{aligned} \quad (70)$$

$$= K \frac{\mathbb{E}_P[L^2]}{\mathbb{E}_P[L]} + \mathbb{E}_{\tilde{P}}[L]. \quad (71)$$

APPENDIX D PROOF OF THEOREM 3

Assume at the optimal θ , we have $\sum_{i=1}^2 2^{-l_i} < \frac{1}{2}$ and $\mathbb{E}_P[L] > \frac{1}{Df_{max}}$. Then, from (14) and (15), we have $\gamma = \lambda = 0$. Then, from (13), we have,

$$2Kl_i p_i + \tilde{p}_i \mathbb{E}_P[L] + p_i \mathbb{E}_{\tilde{P}}[L] = \theta p_i. \quad (72)$$

Then, summing both sides with respect to i , we have,

$$(2K+1)\mathbb{E}_P[L] + \mathbb{E}_{\tilde{P}}[L] = \theta. \quad (73)$$

Then, from (72), we can obtain that,

$$l_i = \left(2K + 1 - \frac{\tilde{p}_i}{p_i} \right) \frac{\mathbb{E}_P[L]}{2K} \quad (74)$$

$$= \left(1 + \frac{p_i - \tilde{p}_i}{2Kp_i} \right) \mathbb{E}_P[L]. \quad (75)$$