Learning the Performances of Intelligent Vehicles with Gaussian Mixture and Monotonicity Information

Zhiyuan Huang
Department of Industrial and
Operations Engineering
University of Michigan

Ding Zhao
University of Michigan
Transportation Research Institude
zhaoding@umich.edu

Henry Lam
Department of Industrial and
Operations Engineering
University of Michigan

Abstract—This paper proposes a new framework based on Gaussian Mixture Model for evaluating risks of Automated Vehicles in a naturalistic driving environment. We extended the previous studies on the Accelerated Evaluation for Automated Vehicles. The proposed procedure includes Gaussian Mixture Model fitting and Importance Sampling distribution construction for Gaussian Mixture Model. The Importance Sampling exploits the monotonic property of the rare event of interest. We present simulation results to confirm that the procedure is valid.

I. Introduction

As Automated Vehicles (AVs) entering the market, it is important to form a systematic evaluation procedure for safety testing. The Naturalistic Field Operational Test (N-FOT) [1], which is adopted by many companies, is inefficient due to the rareness of safety critical scenarios.

In [2], we proposed a brand new concept, the Accelerated Evaluation, and used the procedure to evaluate Automated Vehicles crash risks in natural driving environment. We mainly explored the application on models of interaction between AVs and human-controlled vehicles. In [3], [4], we modeled the frontal crash with a lead vehicle using single variable Gaussian process. We also explored different approaches to tackle the evaluation of lane change scenario [2], [5].

In previous work, our studies on the lance change scenario are based on the independence of random variables. We managed to decompose the density distribution and model the randomness using single variate distributions. In [2], we used single parametric distribution to model the variables and we proposed the piecewise mixture distribution [6] to increase the accuracy of model and evaluation efficiency. Cross Entropy method was used to optimize the importance sampling distribution. While the independence of variables is concluded from data observation, it is not strictly proved. Ignoring the dependence between variables might lead to some estimation error. Moreover, the independence of variables is not a general condition. The requirement of independence restraints the application domain of the Accelerated Evaluation procedure.

In this paper, we proposed an Accelerated Evaluation procedure using Gaussian Mixture Model. Gaussian Mixture Model is suitable and flexible for multivariate data. Data fitting using the Gaussian Mixture Model not only enabled us to capture the dependency between variables, but also provide accuracy if the number of mixture components is appropriate.

Due to the numbers of parameters, using the Cross Entropy method to construct Importance Sampling distribution can be inefficient. The Cross Entropy method requires to solve a stochastic optimization problem for several iterations. When the number of parameters is large, we need more samples to obtain a reasonable solution. This largely increases the required number of samples in the overall procedure. To avoid the inefficiency of the Cross Entropy method, We develop a new procedure based on the monotonicity property of rare event set, which we define in this paper. We derive a rare event set learning procedure to construct Importance Sampling distribution using both monotonicity property of rare-event sets and efficiency of change of measure for Gaussian distributions.

Section II reviews the setting of the lane change scenario. In Section III, we review the truncated Gaussian Mixture Model fitting. The definition of monotonicity rare event sets and the set learning algorithm are present in Section V. We introduce how to construct an efficient Important Sampling distribution for Gaussian Mixture Model in Section IV and show the specific procedure for monotonic rare events in VI. We present the simulation results for the lane change scenario in Section VII. Section VIII concludes the paper.

II. The Lane Change Scenario

The scenario studied in this paper is the defined as the following: a human-controlled vehicle driving in front of an automated vehicle start to cut in the lane where the automated vehicle is running. We want to evaluate the risk of crash in this scenario. Since the scenario is initialized by human, we can take it as a random environment. Then we can use the Automated Vehicle model to simulate the output (crash or not) for a given initial condition.

To model the initial condition of lane changes, we extracted data from the Safety Pilot Model Deployment (SPMD) database [7]. The database includes over 2 million miles of vehicle driving data collected from 98 cars over 3 years. we identify 403,581 lane change events and use 173,692 events
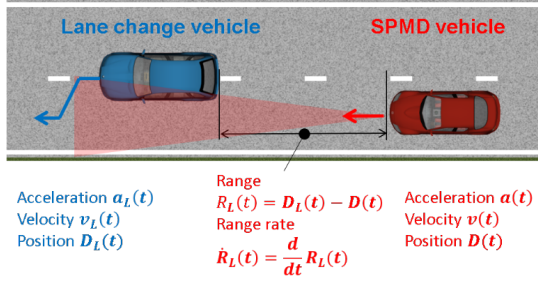
Fig. 1. Lane change data collected by SPMD vehicle.

with a negative range rate to build statistical models. Three key variables can capture the effects of gap acceptance of the lane changing vehicle: velocity of the lead vehicle v, range to the lead vehicle R and time to collision TTC. TTC was defined as:

$$TTC = -\frac{R}{\dot{R}}, \tag{1}$$

where $\dot{R}$ is the relative speed.

The simulation of Automated Vehicle model is based on Adaptive Cruise Control (ACC) and Autonomous Emergency Braking (AEB) [8] regarding the surrounding environment. With an initial condition, the simulator returns an output of the scenario. We consider as an event indicator function $I_\varepsilon(x)$ that returns 1 (crash) or 0 (safe) depending on the event of interest.

III. GAUSSIAN MIXTURE MODEL FITTING

In this section, we review the truncated Gaussian Mixture Model (GMM) fitting studied by Lee and Scott [9].

The fitting of Gaussian Mixture Model is well studied and generally used. In the lane change scenario, the data need to be considered as truncated. For instance, the range of two cars will not be negative. For this reason, we use the truncated GMM to model the initial condition.

For a d dimensional dataset with truncated as a hyper-rectangle with vertices s and t, observations y^n satisfies $s \le y^n \le t$. Note that we can have $s_i = -\infty$ or $t_i = \infty$, which indicates the ith coordinate is not truncated below or above.

For a truncated GMM with K components, we use observation to estimate parameters in

$$g(y) = \sum_{k=1}^{K} \eta_k g_k(y), \tag{2}$$

where η_k is the mixing weights, $\sum_{k=1}^{K} \eta_k = 1$, g_k is truncated Gaussian with support $[s, t]$, mean μ_k and covariance Σ_k. Similarly as the vanilla GMM fitting, we use an Expectation-Maximization (EM) algorithm to estimate the parameters. With a proper initial value for the estimated parameters, we use the following algorithm to iterate for converged estimators.

The E-step is:

$$\langle z_k^n \rangle = \frac{\eta_k g_k(y^n)}{\sum_l \eta_l g_l(y^n)}, \tag{3}$$

where we define $\langle z_k^n \rangle = P(z_k^n = 1|y^n)$ to denote the probability that y^n is generated from the kth component.

The M-step is also similar to the ordinary GMM fitting except some correction terms:

$$\hat{\eta}_k = \frac{1}{N} \sum_n \langle z_k^n \rangle, \tag{4}$$

$$\hat{\mu}_k = \frac{\sum_n \langle z_k^n \rangle y^n}{\sum_n \langle z_k^n \rangle} - m_k, \tag{5}$$

$$\hat{\Sigma}_k = \frac{\sum_n \langle z_k^n \rangle (y^n - \hat{\mu}_k)(y^n - \hat{\mu}_k)^T}{\sum_n \langle z_k^n \rangle} + H_k, \tag{6}$$

where

$$m_k = \mathcal{M}^1(0, \Sigma_k; [s - \mu_k, t - \mu_k]), \tag{7}$$

$$H_k = \Sigma_k - \mathcal{M}^2(0, \Sigma_k; [s - \mu_k, t - \mu_k]). \tag{8}$$

$\mathcal{M}^1(\mu, \Sigma; [a, b])$ and $\mathcal{M}^2(\mu, \Sigma; [a, b])$ denote the first and second moment of truncated Gaussian distribution with truncating vertices a, b, mean μ and covariance Σ.

The choice of number of components K can be determined by some criteria for goodness of fitting, for example, we can use the Bayesian Information Criterion (BIC).

IV. IMPORTANCE SAMPLING DISTRIBUTION FOR GAUSSIAN MIXTURE MODELS

In this section, we first review the concept of Importance Sampling and some known Importance Sampling schemes for rare events with Gaussian distribution. We derive the scheme for GMM based on these background knowledge.

A. Importance Sampling for Gaussian Distribution

Importance Sampling is a technique to reduce the variance in simulation.

Consider a random vector x with distribution F and a rare event set $\varepsilon \subset \Omega$ on sample space Ω. Our goal is to estimate the probability of the rare event

$$P(X \in \varepsilon) = E[I_\varepsilon(X)] = \int I_\varepsilon(x) dF, \tag{9}$$

where the event indicator function is defined as

$$I_\varepsilon(x) = \begin{cases} 1 & x \in \varepsilon. \\ 0 & otherwise. \end{cases} \tag{10}$$

The crude Monte Carlo using the sample mean of $I_\varepsilon(x)$

$$\hat{P}(X \in \varepsilon) = \frac{1}{N} \sum_{n=1}^{N} I_\varepsilon(X_n), \tag{11}$$

where X_i's are drawn from distribution F.

The Importance Sampling [10] technique is derived from

$$E[I_\varepsilon(X)] = \int I_\varepsilon(x) dF = \int I_\varepsilon(x) \frac{dF}{dF^*} dF^*, \tag{12}$$

which gives the estimator using the sample mean of the above expectation (use ref)

$$\hat{P}(X \in \varepsilon) = \frac{1}{N} \sum_{n=1}^{N} I_\varepsilon(X_n) \frac{dF}{dF^*}, \tag{13}$$

where X_i's are generated from F^*, which has the same support with F. We note that this is an unbiased estimation of $P(X \in \varepsilon)$. By appropriately selecting F^*, the evaluation procedure obtains an estimation with smaller variance. F^* is called the IS distribution. We define the likelihood function

$$L(x) = \frac{dF(x)}{dF^*(x)}. \tag{14}$$

The analysis of asymptotic efficiency [10], [11] is a benchmark to determine whether an IS distribution is proper. Let $Z = L(x)I_\varepsilon(x)$ be an IS estimator, it is efficient if

$$\lim_{\varepsilon \to \infty} \frac{\log \left(E_{F^*}[Z^2] \right)}{\log \left(E_{F^*}[Z] \right)} \leq 2, \tag{15}$$

where $\varepsilon \to \infty$ denotes that the rare event set ε diverges to ∞ in a suitable sense (e.g., $\inf_{x \in \varepsilon} \|x\|_2 \to \infty$).

When x follows Gaussian distribution with mean μ and covariance matrix Σ and the rare event set ε satisfies certain assumptions, there is a simple scheme that obtains an efficient IS distribution. Here, we introduce the scheme with assumptions satisfied in the lane change scenario.

For a convex set rare event set ε, we define the dominating point of ε on $\phi(\cdot; \mu, \Sigma)$ to be

$$a^* = \arg \max_{a \in \varepsilon} \phi(a; \mu, \Sigma), \tag{16}$$

where $\phi(\cdot; \mu, \Sigma)$ is the density function for Gaussian distribution with mean μ and covariance matrix Σ. The dominating point contributes the highest density among all points in ε. By shifting the mean μ of the Gaussian distribution to a^*, we can obtain an IS distribution that provides a efficient estimator for $P(x \in \varepsilon)$ [12], [13].

B. IS Scheme for Gaussian Mixture Model and Union of Convex Rare Event Sets

Based on the IS scheme for Guassian distribution on convex rare event set, we propose an IS scheme for Gaussian Mixture Model on the union of convex rare event sets.

Assume x follows a Gaussian Mixture Model with k components, the density of x is

$$f(x) = \sum_{i=1}^{k} p_i \phi(x; \mu_i, \Sigma_i). \tag{17}$$

The rare event set ε is consisted with l convex sets, we denote as $\varepsilon = \cup_{j=1}^{l} \varepsilon_j$.

For each convex set ε_j, we find the dominating point of ε_j on the Gaussian component $\phi(x; \mu_i, \Sigma_i)$ by

$$a_{ij}^* = \arg \max_{a \in \varepsilon_j} \phi(a; \mu_i, \Sigma_i). \tag{18}$$

We propose to use the IS distribution as following:

$$f^*(x) = \sum_{i=1}^{k} \sum_{j=1}^{l} p_i q_j \phi(x; a_{ij}, \Sigma_i), \tag{19}$$

where the q_j can be arbitrary positive number that satisfies $\sum_{j=1}^{l} q_j = 1$. We use $q_j = 1/l$.
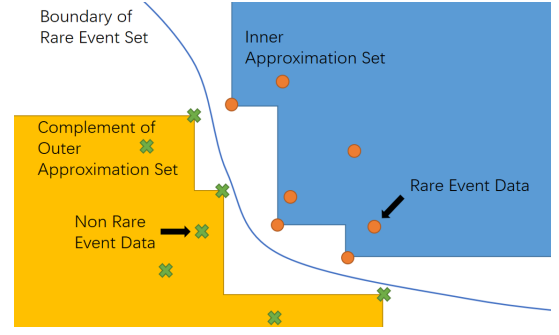

Fig. 2. An illustration of the monotonic set learning.

V. MONOTONIC RARE EVENT SETS LEARNING

In this section, we first define the monotonicity for rare event sets. Then we propose a learning algorithm to obtain an outer approximation set, which contains the rare event set, and an inner approximation set, which is a subset of the rare event set, based on the monotonicity property.

A. Definition of Monotonic Rare Event Sets

For a rare event set ε on d dimensional space, if $x_1 \in \varepsilon$ and $x_1 \leq x_2$ implies that $x_2 \in \varepsilon$, we define the set ε non-decreasing. We note that non-increasing set can be defined by flipping inequality as $x_1 \geq x_2$. Both non-decreasing and non-increasing set are defined to be monotonic.

For example, in the lane change scenario, if a crash occurs for initial condition (v, TTC, R), then if any of these variable is smaller, we can determine that a crash will happen. The set for crash is non-increasing. This can be explained intuitively, because a smaller v, TTC and R means that there is less room for the following Automated Vehicle to make adjustment.

B. Learning Algorithm

Since our goal for learning the approximation set of rare event is to construct an IS distribution, we want the approximated set to satisfy the assumption in the scheme we proposed in Section IV. Therefore, we want the approximation set to be an union of convex sets.

Let us consider a non-decreasing rare event set ε on a d dimensional space. We denote $\mathcal{S}_1 = \{a^1, ..., a^{n_1}\}$ as the observed data in the rare event set and $\mathcal{S}_0 = \{b^1, ..., b^{n_0}\}$ as the observed data not contained by the rare event set. We use $\mathcal{S}_1$ to construct an inner approximation set $\underline{\varepsilon}$ and use $\mathcal{S}_0$ to construct an outer approximation set $\bar{\varepsilon}$.

For each data point a_i in $\mathcal{S}_1$, we construct a set $\mathcal{I}_i = \{x : x \geq a_i\}$. By the definition of non-decreasing set, we have $\mathcal{I}_i \subset \epsilon$. Therefore, we have $\underline{\varepsilon} = \cup_{i=1}^{n_1} \mathcal{I}_i \subset \epsilon$ as the inner approximation set of ε.

For each data point b_j in $\mathcal{S}_0$, we construct sets $\mathcal{O}_{jk} = \{x : x^k \geq b_j^k\}$ for $k = 1, ..., d$, where x^k denotes the kth element in x and b_j^k denotes the kth element in b_j. We have $\epsilon \subset \cup_{k=1}^{d} \mathcal{O}_{jk}$, which indicates that $\epsilon \subset \cup_{j=1}^{n_0} \cup_{k=1}^{d} \mathcal{O}_{jk}$. We define $\bar{\varepsilon} = \cup_{j=1}^{n_0} \cup_{k=1}^{d} \mathcal{O}_{jk}$ as the outer approximation set of ε. Fig. 2 illustrates the key idea of this learning procedure.

We note that $\mathcal{I}_i$'s and $\mathcal{O}_{jk}$'s are convex, the approximation sets are unions of convex sets. We also note that using the minimums of $\mathcal{S}_1$ or maximums of $\mathcal{S}_0$ (also known as Pareto fronts) of the dataset provides the same inner or outer approximation set. Here, minimums are defined as points that are greater in all dimensions than no other points; whereas maximums are defined as points that are smaller in all dimensions than no other points.

Additionally, to use the result for non-decreasing sets, we can simply flip all coordinates of data (take negative value) when we have a non-increasing set. For sets that is non-decreasing on some of the coordinates and non-increasing on the rest coordinates, we can flip the value of data those non-increasing coordinates. The approximation sets obtained will still be union of convex sets after flipping any coordinate.

We note that we can directly obtain a lower probability bound and an upper probability bound for the rare event probability.

VI. IS DISTRIBUTION CONSTRUCTION SCHEME FOR A GAUSSIAN MIXTURE MODEL OF A MONOTONIC RARE EVENT SET

Combining the discussion in Sections IV and V, we propose an iterative procedure that provides IS distributions for a Gaussian Mixture Model with a monotonic rare event set.

Here, we consider a non-decreasing rare event set ε on d dimensional space and the variable vector x is generated from a k components GMM with density (17). To clarify the notations, for each Gaussian component i, we construct a set of dominating points $\mathcal{A}_I^i$ using the inner approximation set $\underline{\varepsilon}$ and $\mathcal{A}_O^i$ using the outer approximation set $\bar{\varepsilon}$ for $i = 1, ..., k$. $|\mathcal{A}|$ denotes the number of elements in the set $\mathcal{A}$. $\underline{\varepsilon}$ and $\bar{\varepsilon}$ are constructed based on $\mathcal{S}_1$, which contains the minimums of observed data in ε, and $\mathcal{S}_0$, which contains the maximums of observed data that not in ε. The procedure iterates to update these sets, which will provide two IS distributions f_I^* and f_O^* based on the inner approximation and the outer approximation of ε respectively.

The procedure is presented as the following:

1) Initialize $A_I^i = A_O^i = \{\mu_i\}$ and $\mathcal{S}_1 = \mathcal{S}_0 = \emptyset$.
2) Construct the sampling distribution

$$f_I^*(x) = \sum_{i=1}^{k} \sum_{a \in \mathcal{A}_I^i} p_i \frac{1}{|\mathcal{A}_I^i|} \phi(x; a, \Sigma_i) \qquad (20)$$

and

$$f_O^*(x) = \sum_{i=1}^{k} \sum_{a \in \mathcal{A}_O^i} p_i \frac{1}{|\mathcal{A}_O^i|} \phi(x; a, \Sigma_i) \qquad (21)$$

3) Sample N data points $D = \{x_1, ..., x_N\}$ from the density

$$f(x) = \rho f_I^*(x) + (1 - \rho) f_O^*(x), \qquad (22)$$

where ρ is on $[0, 1]$. We suggest to use $\rho = 1/2 \, I_{\{\mathcal{S}_1 \neq \{\mu_i\}\}}$.

4) Input D to the simulator $I_\varepsilon(x)$ and use the outcome to update S_1 and S_0. We add the new data points to S_1 and S_0 regarding the outcome of $I_\varepsilon(x)$. Then we discard non-minimum data points in S_1 and non-maximum data points in S_0.
5) For each Gaussian component i, we use each data points $a \in S_1$ to solve

$$\max_x \ \phi(x; \mu_i, \Sigma_i) \text{ subject to } x \geq a. \qquad (23)$$

We obtain $|S_1|$ solutions and the solution set is our new $\mathcal{A}_I^i$.
6) For each Gaussian component i, we use each data points $b \in S_0$ to solve

$$\max_x \ \phi(x; \mu_i, \Sigma_i) \text{ subject to } x_m \geq b_m, \qquad (24)$$

for $m = 1, ..., d$. $x_m \geq b_m$ denotes the mth element of x is greater or equals to the mth element of b. We obtain $d|S_0|$ solutions and the solution set is our new $\mathcal{A}_O^i$.
7) Iterate from 2) to 6).

The stop criterion can either be a maximum iteration number or a maximum number of elements in certain set if reached. A recommended IS distribution is in the form of (22). Since the optimal ρ is case-by-case, we need to observe the elements in $\mathcal{A}_I^i$'s and $\mathcal{A}_O^i$'s to decide the value of ρ. If there elements in $\mathcal{A}_I^i$'s and $\mathcal{A}_O^i$'s are very different (in the sense of variable values) or there are few elements in , we recommend to use $\rho = 0$. If $\mathcal{A}_I^i$'s and $\mathcal{A}_O^i$'s are similar, then we use $\rho = 0.5$.

VII. SIMULATION ANALYSIS ON LANE CHANGE SCENARIO

In this section, we present the GMM fitting of the lane change model and use the IS distribution construction scheme proposed on the model. We show simulation results to justify the validity of the scheme.

A. Truncated Gaussian Mixture Model Fitting

We use the EM algorithm in Section III to fit the lane change data (v, TTC^{-1}, R^{-1}). The important part for GMM fitting is the selection of the number of components K. Fig. 3 shows the BIC regards to different K. We could observe that we obtain a local minimum at $K = 9$, where it means that the model with $K = 9$ provides a balance between the number of parameters and the fitting.

We note that the different scale of the variables might cause some numerical issues in implementing the algorithm presented in Section III, we can normalize (subtract by marginal mean and then divided by marginal standard deviation) the data before we fit the model.

B. Importance Sampling Results

The IS distribution construction scheme we proposed in this paper is based on the assumption of ordinary Gaussian Mixture Model. Since the distribution of variables in the lane change scenario is truncated, we need to make small modifications on the scheme.
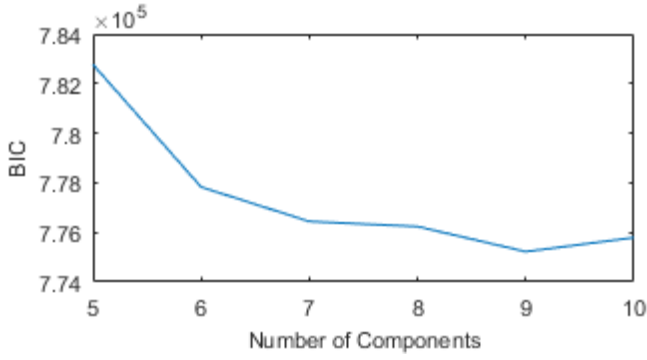
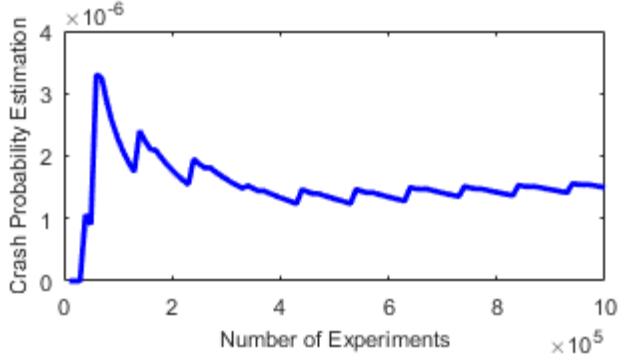Fig. 3. The BIC regarding to different number of components K.

Fig. 5. The confidence interval of the crash probability estimation with increasing number of samples.

Fig. 4. The crash probability estimation with increasing number of samples.

algorithm increases the efficiency of the estimation by roughly 25 times.

VIII. CONCLUSION

This paper propose using Gaussian Mixture Model to model stochastic variables in Automated Vehicle evaluating problems. We provide a algorithm for constructing Importance Sampling distribution based on the property of Gaussian Mixture model and monotonic rare events. The proposed algorithm can provide a valid Importance Sampling distribution.

A further direction for our research is to find a scheme to refine the elements in the dominating point sets.

Shifting the sampling mean to dominating points still works in for truncated Gaussian, because the truncated coordinate does not have asymptotic behavior. In this case, we only need to worry about those untruncated coordinates, which will be the same as untruncated Gaussian Distribution.

For the monotonic set learning part, note that if we directly use (23) and (24) for truncated variables, we might obtain in the infeasible region. We add the truncated boundary as constrains for these optimization problems.

After we obtain the dominating point sets $\mathcal{A}_I^i$'s and $\mathcal{A}_O^i$'s from the proposed algorithm, we observe that the elements in $\mathcal{A}_I^i$ and $\mathcal{A}_O^i$ are tend to be very different. Therefore, we use $\rho = 0$ for the IS distribution.

Fig. 4 presents the estimated probability with different number of experiments. The probability converges around 1.15×10^6. Fig. 5 shows the 95% confidence interval half width of the probability estimation. For 10^6 samples, we have the 95% confidence interval for the estimation as $(1.02 \times 10^6, 1.97 \times 10^{-6})$. Using the formula for the standard deviation of the crude Monte Carlo estimation $\hat{P}(x \in \varepsilon)$ by

$$std(\hat{P}(x \in \varepsilon)) = \sqrt{\frac{\hat{P}(x \in \varepsilon)(1 - \hat{P}(x \in \varepsilon))}{n}}, \qquad (25)$$

we can estimate that the crude Monte Carlo method requires about 2.56×10^7 samples to reach a confidence interval with a similar scale. The IS distribution constructed by the proposed
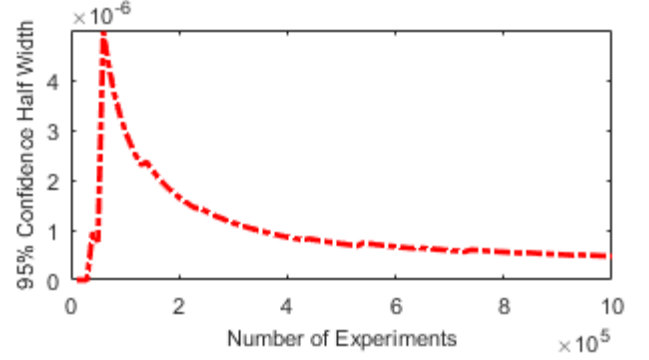
REFERENCES

[1] FESTA-Consortium, "FESTA Handbook Version 2 Deliverable T6.4 of the Field opErational teSt supporT Action," FESTA, Tech. Rep., 2008.

[2] D. Zhao, H. Lam, H. Peng, S. Bao, D. J. LeBlanc, K. Nobukawa, and C. S. Pan, "Accelerated Evaluation of Automated Vehicles Safety in Lane-Change Scenarios Based on Importance Sampling Techniques," *IEEE Transactions on Intelligent Transportation Systems*, 2016.

[3] D. Zhao, H. Peng, S. Bao, K. Nobukawa, D. J. LeBlanc, and C. S. Pan, "Accelerated evaluation of automated vehicles using extracted natural-istic driving data," in *Proceeding for 24th International Symposium of Vehicles on Road and Tracks*, 2015.

[4] D. Zhao, X. Huang, H. Peng, H. Lam, and D. J. Leblanc, "Accelerated Evaluation of Automated Vehicles in Car-Following Maneuvers," *ArXiv no. 1607.02687*, 2016.

[5] Z. Huang, D. Zhao, H. Lam, and D. J. LeBlanc, "Accelerated evaluation of automated vehicles using piecewise mixture models," *arXiv preprint arXiv:1701.08915*, 2017.

[6] Z. Huang, D. Zhao, H. Lam, D. J. LeBlanc, and H. Peng, "Evaluation of Automated Vehicles in the Frontal Cut-in Scenario-an Enhanced Approach using Piecewise Mixture Models," 10 2016. [Online]. Available: http://arxiv.org/abs/1610.09450

[7] D. Bezzina and J. R. Sayer, "Safety Pilot: Model Deployment Test Conductor Team Report," NHTSA, Tech. Rep., 2014.

[8] A. G. Ulsoy, H. Peng, and M. Çakmakci, *Automotive control systems*. Cambridge University Press, 2012.

[9] G. Lee and C. Scott, "EM Algorithms for Multivariate Gaussian Mixture Models with Truncated and Censored Data," *Computational Statistics and Data Analysis*, vol. 56, no. 9, pp. 2816–2829, 2012.

[10] J. Bucklew, *Introduction to Rare Event Simulation*. Springer Science & Business Media, 2004.

[11] S. Asmussen and P. Glynn, *Stochastic Simulation: Algorithms and Analysis*.   Springer, 2007.

[12] J. S. Sadowsky *et al.*, "On monte carlo estimation of large deviations probabilities," *The Annals of Applied Probability*, vol. 6, no. 2, pp. 399–422, 1996.

[13] A. Dieker and M. Mandjes, "On asymptotically efficient simulation of large deviation probabilities," *Advances in applied probability*, vol. 37, no. 02, pp. 539–552, 2005.