Uncertainty-Aware Distributionally Robust Model Predictive Control for Safe Autonomous Driving

Hyeongchan Ham, Heejin Ahn
Korea Advanced Institute of Science and Technology
{hyeongchan.ham, heejin.ahn}@kaist.ac.kr

Abstract: Safety is a critical concern in motion planning for autonomous vehicles. Modern autonomous vehicles rely on neural network-based perception, but making control decisions based on these inference results poses significant safety risks due to inherent uncertainties. To address this challenge, we present a distributionally robust optimization (DRO) framework that accounts for both aleatoric and epistemic perception uncertainties using evidential deep learning (EDL). Our approach introduces a novel ambiguity set formulation based on evidential distributions that dynamically adjusts the conservativeness according to perception confidence levels. We integrate this uncertainty-aware constraint into model predictive control (MPC), proposing the DRO-EDL-MPC algorithm with computational tractability for autonomous driving applications. Validation in the CARLA simulator demonstrates that our approach maintains efficiency under high perception confidence while enforcing conservative constraints under low confidence. Videos are available at https://dro-edl-mpc.github.io/

Keywords: robot safety, planning under uncertainty, machine learning for control

1 Introduction

Autonomous driving systems have attracted significant attention due to their potential impacts [1]. These systems rely on three fundamental technologies: perception to interpret the environment, planning to determine optimal trajectories, and control to execute driving actions. One of the critical challenges in developing safe autonomous driving systems lies in the inherent uncertainties of the perception module, which propagate through the decision-making pipeline [2]. These uncertainties stem from noises in the data, known as *data uncertainty* or *aleatoric uncertainty*, and the lack of knowledge during the training procedure, known as *model uncertainty* or *epistemic uncertainty* [3].

To address these uncertainties, researchers have proposed various approaches for safe and robust decision-making in autonomous driving systems. For handling data uncertainty specifically, risk metrics-based approaches have been developed, including chance-constrained optimization [4], [5] and conditional value-at-risk (CVaR) methods [6]. While effective in principle, these approaches assume that perception models accurately represent data uncertainty distributions. In practice, perception models suffer from inherent model uncertainty arising from limited training data, architectural constraints, and optimization challenges, leading to incomplete knowledge representations [7].

Distributionally robust optimization (DRO) has emerged as a promising framework to account for such model uncertainty by optimizing against worst-case distributions within a predefined set of candidate distributions, called an ambiguity set [8]. Ambiguity set formulations can be categorized into two main approaches: discrepancy-based [9]–[12] and moment-based [13], [14]. Discrepancy-based approaches define an ambiguity set as a set of distributions within fixed radii in terms of discrepancy metrics, such as ϕ-divergence [9], [10] or Wasserstein distance [11], [12]. These approaches often lead to overly conservative results even when the predicted distribution shows high confidence. Moment-based approaches leverage sample quantities to establish confidence intervals,

Workshop on Safe and Robust Robot Learning for Operation in the Real World at CoRL 2025.

thereby reducing the size of ambiguity sets as the number of samples increases [13]. However, the adaptability of the ambiguity set requires a large number of samples, which is often impractical and undesirable in real-world scenarios.

Recent work in [14] addresses this sample requirement limitation through the application of evidential deep learning (EDL) [15]. Unlike traditional deep learning methods that directly output target predictions, EDL takes a probabilistic approach by learning the parameters of a posterior distribution, referred to as the *evidential distribution*. This distribution represents uncertainty over the parameters of the likelihood distribution of the target. However, the method in [14] leverages the variance of the likelihood variance, which limits its applicability to confident predictions, as the variance becomes undefined for highly uncertain predictions.

We propose a principled framework for constructing ambiguity sets that leverages this higher-order evidential distribution, which addresses the limitation of previous EDL-based approaches by directly integrating the evidential distribution. Our approach defines ambiguity sets by establishing boundaries where the cumulative probability of the evidential distribution reaches a predefined threshold. This construction yields ambiguity sets with clear probabilistic semantics: each set contains the true uncertainty distribution of the data with a predetermined confidence level. Unlike discrepancy-based methods that employ static distance metrics irrespective of model confidence, or moment-based approaches that only constrain particular statistical moments, our method dynamically adjusts the ambiguity set size according to the model uncertainty. Furthermore, in contrast to previous work [14], our approach is broadly applicable to uncertain scenarios.

We implement this EDL-based ambiguity set formulation within the distributionally robust safety constraint, called DR-EDL-CVaR, and this constraint is employed to our proposed DRO-EDL-MPC algorithm for autonomous driving. This algorithm adapts conservativeness based on perception confidence while maintaining computational tractability.

Our main contributions are as follows:

- We propose DR-EDL-CVaR, a distributionally robust safety constraint leveraging the cumulative probability of the evidential distribution to construct an uncertainty-aware ambiguity set. This provides an informative representation of the uncertainties.
- We introduce DRO-EDL-MPC, a computationally tractable motion planning algorithm that incorporates DR-EDL-CVaR.
- We validate DRO-EDL-MPC in the CARLA simulator, demonstrating its effectiveness in enhancing safety under various perception uncertainty conditions.

2 Problem Formulation

The goal of this paper is to compute the optimal motion of an ego vehicle navigating an environment with static obstacles while explicitly accounting for and adapting to perception uncertainties.

The dynamics of the ego vehicle consists of the nominal dynamics model $f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_x}$ and unknown dynamics $g : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_x}$, represented as

$$\mathbf{x}(t+1) = f(\mathbf{x}(t), \mathbf{u}(t)) + g(\mathbf{x}(t), \mathbf{u}(t)), \tag{1}$$

where $\mathbf{x}(t) \in \mathbb{X} \subseteq \mathbb{R}^{n_x}$ and $\mathbf{u}(t) \in \mathbb{U} \subseteq \mathbb{R}^{n_u}$ are the state of the ego vehicle and the control input at timestep t. The nominal dynamics can be derived from physical principles of the system, while the unknown components can be learned from data using Gaussian Process Regression (GPR) [16].

The ego state $\mathbf{x}$ consists of the center positions $\mathbf{c}^x \in \mathbb{R}^{n_c}$, the heading angle $\phi^x \in \mathbb{R}$, and the speed $v^x \in \mathbb{R}$ of the ego vehicle. That is, $\mathbf{x} = (\mathbf{c}^x, \phi^x, v^x)$. For a 2-dimensional example, $n_c = 2$ and $\mathbf{c}^x = (c_1^x, c_2^x)$ represents the coordinates of the center of the ego vehicle. Also, the ego has attributes for the lengths from the center to the boundary $\mathbf{b}^x \in \mathbb{R}^{n_c}$, where, for example, $\mathbf{b}^x = (b_1^x, b_2^x)$ represents the half of the width and length of the vehicle.

The ego vehicle must avoid collisions with static obstacles in the environment. Similar to the ego state $\mathbf{x}$, the obstacle state $\xi \in \mathbb{R}^{n_\xi}$ consists of the center positions $\mathbf{c}^\xi \in \mathbb{R}^{n_c}$, the heading angle $\phi^\xi \in \mathbb{R}$, and the speed $v^\xi \in \mathbb{R}$. That is, $\xi = (\mathbf{c}^\xi, \phi^\xi, v^\xi)$. The obstacle also has the attributes for the lengths from the center to the boundary $\mathbf{b}^\xi \in \mathbb{R}^{n_c}$. To prevent collisions between the ego vehicle and obstacles, we define a safety loss function $\ell : \mathbb{R}^{n_x} \times \mathbb{R}^{n_\xi} \to \mathbb{R}$ to quantify the level of safety. The safety constraint is encoded by

$$\ell(\mathbf{x}, \xi) = (r^x + r^\xi)^2 - \|\mathbf{c}^x - \mathbf{c}^\xi\|_2^2 \leq 0, \tag{2}$$

where $r^\mathbf{s}$ denotes the radius of any state $\mathbf{s} \in \{x, \xi\}$ and is defined as $r^\mathbf{s} := \|\mathbf{b}^\mathbf{s}\|_2 = \sqrt{\sum_i^{n_c} (b_i^\mathbf{s})^2}$.

The obstacle state ξ is estimated from sensor data o using a perception model F during the initialization stage. Since these prediction results inherently contain uncertainties, we must account for both data and model uncertainties. To address data uncertainty, we employ Conditional Value-at-Risk (CVaR) constraint, $\text{CVaR}_\epsilon^{\mathbb{P}}[\ell(\mathbf{x}, \xi)] \leq 0$[1], which enables the evaluation of expected loss in extreme scenarios. Here, $\mathbb{P}$ represents the probability distribution of ξ, and $\epsilon \in [0.5, 1)$ is the confidence level parameter. For model uncertainty, we use a distributionally robust approach. Given an ambiguity set $\mathbb{D}$ containing distribution $\mathbb{P}$, we formulate the distributionally robust CVaR constraint as $\max_{\mathbb{P}\in\mathbb{D}} \text{CVaR}_\epsilon^{\mathbb{P}}[\ell(\mathbf{x}, \xi)] \leq 0$. This constraint ensures safety by optimizing against the worst-case distribution within the ambiguity set.

The distributionally robust MPC problem with the stage-wise cost c and the terminal cost q is formulated as follows:

$$\min_{\mathbf{u}} \quad \sum_{t=0}^{T-1} c(\mathbf{x}(t), \mathbf{u}(t)) + q(\mathbf{x}(T)) \tag{3a}$$

$$\text{s.t.} \quad \mathbf{x}(t) \in \mathbb{X},\ \forall t \in \mathbb{Z}_{0:T},\ \ \mathbf{u}(t) \in \mathbb{U},\ \forall t \in \mathbb{Z}_{0:T-1} \tag{3b}$$

$$(1),\ \forall t \in \mathbb{Z}_{0:T-1} \tag{3c}$$

$$\max_{\mathbb{P}\in\mathbb{D}} \text{CVaR}_\epsilon^{\mathbb{P}}[\ell(\mathbf{x}(t), \xi)] \leq 0,\ \forall t \in \mathbb{Z}_{0:T}. \tag{3d}$$

Here, $\mathbb{Z}_{0:T}$ denotes the set of integers $\{0, 1, \ldots, T\}$, and T is the length of the planning horizon. The obstacle ξ is assumed to be constant during the MPC prediction steps as it represents a static obstacle.

3 EDL-Based Safety Constraint

EDL assumes that the regression target c_i^ξ is a random variable drawn from a Gaussian distribution $\mathcal{N}(\mu_i, \sigma_i^2)$ with the mean μ_i and variance σ_i^2. Also, it assumes that the mean μ_i follows a Gaussian prior $\mathcal{N}(\gamma_i, \sigma_i^2/\lambda_i)$ with $\gamma_i \in \mathbb{R}, \lambda_i > 0$ and the variance σ_i^2 follows the Inverse-Gamma prior with $\alpha_i > 1, \beta_i > 0$. Let us denote $\theta = (\mu_i, \sigma_i^2)$ and $m_i = (\gamma_i, \lambda_i, \alpha_i, \beta_i)$. Then the posterior $p(\theta|m_i)$ is defined as the Normal Inverse-Gamma (NIG) distribution $NIG(\theta|m_i)$, which is also called the evidential distribution [15]. The trained EDL model estimates the parameter m_i of the evidential distribution, from which the prediction value of c_i^ξ, data uncertainty, and model uncertainty can be computed, respectively, as

$$\mathbb{E}[\mu_i] = \gamma_i,\ \ \mathbb{E}[\sigma_i^2] = \frac{\beta_i}{\alpha_i - 1},\ \ Var[\mu_i] = \frac{\beta_i}{\lambda_i(\alpha_i - 1)}.$$

For an EDL model with the regression target $\mathbf{c}^\xi = (c_1^\xi, \ldots, c_{n_c}^\xi)$, it predicts the parameter $\mathbf{m} = (m_1, \ldots, m_{n_c})$. Because current EDL models, such as [17], predict each m_i separately at each head, we also assume each c_i^ξ is independent and $\mathbf{c}^\xi$ is drawn from an n_c-dimensional multivariate Gaussian distribution $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$ where $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_{n_c})$ and Σ is a diagonal matrix of $(\sigma_1^2, \ldots, \sigma_{n_c}^2)$.

[1] For loss $L \sim \mathbb{P}_L$, $\text{CVaR}_\epsilon^{\mathbb{P}_L}[L] := \min_{z\in\mathbb{R}} \mathbb{E}^{\mathbb{P}_L}\left[z + \frac{(L-z)^+}{1-\epsilon}\right]$ where $(L-z)^+ = \max(L-z, 0)$. This is the average of the worst $(1-\epsilon)\%$ of outcomes.

The evidential distribution can be interpreted as a higher-order distribution that captures model uncertainty, and a lower-order realization of the evidential distribution, which is a normal distribution, represents data uncertainty. We leverage this evidential distribution to construct our ambiguity set and formulate distributionally robust safety constraints, enabling us to account for both model and data uncertainties.

Definition 1. (EDL ambiguity set). *Given a cumulative probability threshold $\eta_i \in \mathbb{R}$ and evidential distribution parameter m_i, the ambiguity set for axis i is defined as*

$$\mathbb{D}_i(\eta_i|m_i) := \left\{ \mathcal{N}(\mu, \sigma^2) | \int_{\theta=(\mu,\sigma^2)} NIG(\theta|m_i) d\theta = \eta_i \right\}. \tag{4}$$

For all axis, given $\eta \in \mathbb{R}$ and $\mathbf{m} = (m_1, \ldots, m_{n_c})$, the ambiguity set $\mathbb{D}(\eta|\mathbf{m})$ is defined as

$$\mathbb{D}(\eta|\mathbf{m}) := \{\mathcal{N}(\boldsymbol{\mu}, \Sigma) | \mathcal{N}(\mu_i, \sigma_i^2) \in \mathbb{D}_i(\sqrt[n_c]{\eta}|m_i), \forall i\}. \tag{5}$$

EDL generates a dispersed evidential distribution under high model uncertainty and a concentrated one when uncertainty is low, causing our ambiguity set to expand or contract accordingly. This ensures that even with a fixed confidence level η_i, our approach automatically adapts its conservativeness based on the perception model's uncertainty.

Next, we formulate the distributionally robust safety loss within this ambiguity set.

Definition 2. (DR-EDL-CVaR). *The distributionally robust safety loss given EDL ambiguity set $\mathbb{D}(\eta|\mathbf{m})$ is defined as*

$$\textit{DR-EDL-CVaR}_\epsilon^{\mathbb{D}(\eta|\mathbf{m})}[\ell(\mathbf{x}, \xi)] := \max_{\mathbb{P} \in \mathbb{D}(\eta|\mathbf{m})} \textit{CVaR}_\epsilon^{\mathbb{P}}[\ell(\mathbf{x}, \xi)]. \tag{6}$$

The safety loss ℓ is defined in (2). This formulation enhances the robustness of our control approach by simultaneously addressing both types of uncertainty: data uncertainty through CVaR risk metric at confidence level ϵ, and model uncertainty by optimizing against the worst-case CVaR within the ambiguity set at confidence level η.

4 Our Proposed Method

4.1 Identification of the worst-case obstacle distribution

Because it is nontrivial to find the closed-form solution of the ambiguity set, we define a *surrogate ambiguity set*, which facilitates calculating the worst-case loss.

Definition 3. (Surrogate ambiguity set). *Let $\Theta := \{\theta : NIG(\theta|m_i) \leq c_{th}\}$ denote a sublevel set that satisfies $\int_\Theta NIG(\theta|m_i)d\theta = \eta_i$. Let $\mu_{i,min}, \mu_{i,max}$ and $\sigma^2_{i,min}, \sigma^2_{i,max}$ be the extreme values of μ and σ^2, respectively, in Θ. Let $\mathcal{I}_{i,\mu} := [\mu_{i,min}, \mu_{i,max}]$ and $\mathcal{I}_{i,\sigma^2} := [\sigma^2_{i,min}, \sigma^2_{i,max}]$. We define the surrogate ambiguity set as*

$$\mathbb{D}_i^{sur}(\eta_i|m_i) := \{\mathcal{N}(\mu, \sigma^2) : \mu \in \mathcal{I}_{i,\mu}, \sigma^2 \in \mathcal{I}_{i,\sigma^2}\}. \tag{7}$$

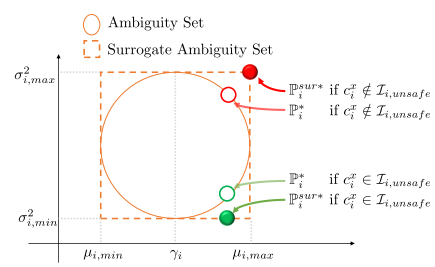

Figure 1: Illustration of moment sets (μ_i, σ_i^2) for the original and surrogate ambiguity sets along axis i. The worst-case distributions within the original and surrogate ambiguity sets are identified as $\mathbb{P}_i^*$ and $\mathbb{P}_i^{sur*}$, respectively.

This means that the surrogate ambiguity set is a set of Gaussian distributions whose μ and σ^2 lie in a rectangular region $\mathcal{I}_{i,\mu} \times \mathcal{I}_{i,\sigma^2}$. The rectangular region $\mathcal{I}_{i,\mu} \times \mathcal{I}_{i,\sigma^2}$ encloses the contour Θ, resulting in for any η_i and m_i

$$\mathbb{D}_i(\eta_i|m_i) \subset \mathbb{D}_i^{sur}(\eta_i|m_i). \tag{8}$$

It also implies that the cumulative probability of the NIG distribution at $\mathcal{I}_{i,\mu} \times \mathcal{I}_{i,\sigma^2}$ is larger than η_i, which makes the ambiguity set more conservative. In Fig. 1, the contour Θ that defines the original ambiguity set and $\mathcal{I}_{i,\mu} \times \mathcal{I}_{i,\sigma^2}$ that defines the surrogate ambiguity set are illustrated as a solid round and a dashed rectangle, respectively.

Proposition 1. (Conservative surrogate ambiguity set solution). *Given the surrogate ambiguity set $\mathbb{D}_i^{sur}(\eta_i|m_i)$, the distributionally robust safety loss satisfies*

$$\max_{\mathbb{P}\in\mathbb{D}(\eta|\mathbf{m})} \textit{CVaR}_\epsilon^{\mathbb{P}}[\ell(\mathbf{x}(t),\xi)] \leq (r^x + r^\xi)^2 - \sum_i^{n_c} \max_{\mathbb{P}_i\in\mathbb{D}_i^{sur}(\eta_i|m_i)} \textit{CVaR}_\epsilon^{\mathbb{P}_i}[-|c_i^x - c_i^\xi|]^2.$$

Furthermore,

- *If $c_i^x \in \mathcal{I}_{i,\mu}$,*
$$\max_{\mathbb{P}_i\in\mathbb{D}_i^{sur}(\eta_i|m_i)} \textit{CVaR}_\epsilon^{\mathbb{P}_i}[-|c_i^x - c_i^\xi|] = \kappa \cdot \sigma_{i,min} < 0,$$
where $\kappa = \frac{1}{1-\epsilon}\sqrt{\frac{2}{\pi}}(\exp(-[erf^{-1}(\epsilon-1)]^2) - 1)$ and $erf^{-1}(\cdot)$ the inverse of the error function.
- *If $c_i^x \in \mathcal{I}_{i,unsafe} \setminus \mathcal{I}_{i,\mu}$,*
$$\max_{\mathbb{P}_i\in\mathbb{D}_i^{sur}(\eta_i|m_i)} \textit{CVaR}_\epsilon^{\mathbb{P}_i}[-|c_i^x - c_i^\xi|] \leq \kappa \cdot \sigma_{i,min} < 0,$$
where $\mathcal{I}_{i,unsafe} = [\mu_{i,min} - \delta \cdot \sigma_{i,max}, \mu_{i,max} + \delta \cdot \sigma_{i,max}]$ for $\delta = \frac{1}{1-\epsilon}\frac{1}{\sqrt{2\pi}}\exp(-[erf^{-1}(2\epsilon-1)]^2) > 0$.
- *If $c_i^x \notin \mathcal{I}_{i,unsafe}$,*
$$\max_{\mathbb{P}_i\in\mathbb{D}_i^{sur}(\eta_i|m_i)} \textit{CVaR}_\epsilon^{\mathbb{P}_i}[-|c_i^x - c_i^\xi|] < -|c_i^x - \gamma_i| + \frac{\mu_{i,max} - \mu_{i,min}}{2} + \delta \cdot \sigma_{i,max} < 0,$$
where γ_i is given in the prediction result m_i.

According to Proposition 1, the worst-case obstacle distribution varies depending on the ego vehicle position c_i^x. A one-dimensional example is illustrated in Fig. 1. When the ego vehicle is distant from the obstacle ($c_i^x > \mu_{i,max} + \delta\sigma_{i,max}$), (an upper bound of) the optimal negative distance CVaR is attained at $\mathbb{P}_i^{sur*} = \mathcal{N}(\mu_{i,max}, \sigma_{i,max}^2)$ (filled red circle) within the surrogate ambiguity set. Within the original ambiguity set, $\mathbb{P}_i^*$ (empty red circle) denotes the worst-case obstacle distribution. In contrast, when the ego vehicle is close to the obstacle ($c_i^x \in \mathcal{I}_{i,unsafe}$), (an upper bound of) the maximum negative distance CVaR is attained at $\mathbb{P}_i^{sur*} = \mathcal{N}(c_i^x, \sigma_{i,min}^2)$ (filled green circle) within the surrogate ambiguity set. Within the original ambiguity set, $\mathbb{P}_i^*$ (empty green circle) denotes the worst-case distribution.

Proposition 1 describes how to identify the worst-case distribution $\mathbb{P}_i^{sur*}$ within the surrogate ambiguity set. Thus, we can directly find the worst-case obstacle distribution instead of focusing on the distribution of the loss function.

However, calculating the extrema $\mu_{i,min}, \mu_{i,max}, \sigma_{i,min}$ and $\sigma_{i,max}$ given the confidence threshold η_i remains computationally challenging, as it requires complex numerical optimization problems. To mitigate this issue, we employ standardized solutions. The NIG distribution consists of the normal distribution and the inverse gamma distribution. These distributions have location parameter γ_i and scale parameter σ_i, β_i. Therefore, we can transpose NIG$(\mu_i, \sigma_i^2|\gamma_i, \lambda_i, \alpha_i, \beta_i)$ into NIG$(\mu_{z,i}, \sigma_{z,i}^2|0, 1, \alpha_i, 1)$ using following relation:

$$\mu_i = \mu_{z,i} \cdot \sqrt{\frac{\beta_i}{\lambda_i}} + \gamma_i, \quad \sigma_i = \sigma_{z,i} \cdot \sqrt{\beta_i}. \tag{9}$$

The shape parameter α_i remains unchanged as it is an intrinsic property of the distribution. The contour of the transposed NIG distribution at confidence η_i is calculated using numerical integration

[18] and the extrema $(\mu_{z,i,min}, \mu_{z,i,max}, \sigma^2_{z,i,min}, \sigma^2_{z,i,max})$ are calculated using the Brent's method [19], which is a numerical root finding method. These extrema are computed offline and stored in the lookup table $\mathcal{M}_\alpha$ for different $\alpha \in [1.01, 10.00]$. At execution time, the perception model estimates the NIG distribution parameters m_i, and the extrema are queried from $\mathcal{M}_\alpha$ without computation overhead. These extrema are transformed back to the extrema of the original distribution using (9). Consequently, the standardized approach and Proposition 1 enables determining the worst-case obstacle distribution $\mathbb{P}_i^{sur*}$ in a computationally tractable manner.

4.2 Conservative constraint for tractable MPC (DRO-EDL-MPC)

The safety constraint can be formulated using the worst-case obstacle distribution derived by Proposition 1. However, these constraints vary depending on the ego position c_i^x, for example, whether $c_i^x \in \mathcal{I}_{i,\mu}$, and thus translate into a mixed-integer programming formulation, which can be solved using the Big-M method [20]. To avoid the Big-M approach, which can be computationally expensive and prone to numerical instability, we introduce a conservative reformulation of the worst-case CVaR that is continuous with the ego state, eliminating the need for integer variables.

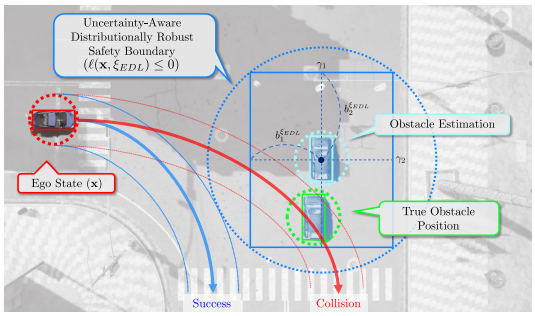

Figure 2: Estimated obstacle position (light blue circle) may differ from the true obstacle position (green circle). By introducing ξ_{EDL} (large blue rectangle), we enforce the constraint $\ell(\mathbf{x}, \xi_{EDL}) \leq 0$ (large blue circle), thereby satisfying the distributionally robust safety constraint.

Proposition 2. (Constraint reformulation). *Given the perception results* $\mathbf{m}, \phi^\xi, v^\xi$, *and* $\mathbf{b}^\xi$, *define* $\xi_{EDL} = (\mathbf{c}^{\xi_{EDL}}, \phi^{\xi_{EDL}}, v^{\xi_{EDL}})$ *as*

$$c_i^{\xi_{EDL}} := \gamma_i, \quad \phi^{\xi_{EDL}} := \phi^\xi, \quad v^{\xi_{EDL}} := v^\xi,$$

and its attribute $\mathbf{b}^{\xi_{EDL}}$ *as*

$$b_i^{\xi_{EDL}} := (\mu_{i,max} - \mu_{i,min})/2 + \delta \cdot \sigma_{i,max} + r^\xi,$$

where δ *is provided in Proposition 1. Then, for any* $\mathbf{x}$,

$$\ell(\mathbf{x}, \xi_{EDL}) \leq 0 \implies \textit{DR-EDL-CVaR}_\epsilon^{\mathbb{D}(\eta|\mathbf{m})}[\ell(\mathbf{x}, \xi)] \leq 0. \tag{10}$$

This proposition defines a new obstacle, ξ_{EDL}, which provides a more conservative constraint that is continuous with respect to c_i^x. Fig. 2 illustrates the perceived obstacle position, true obstacle position and the use of ξ_{EDL}. Distributionally robust and feasible solutions can be found with DRO-EDL-MPC, which integrates (12) into the MPC (3d).

5 Experiments

In this section, we validate the DRO-EDL-MPC algorithm in the CARLA simulator and demonstrate its less conservative behavior under confident perception and more conservative behavior under uncertain environment.

5.1 Experiment Settings

The objective of the ego vehicle is to reach its destination while avoiding collisions with obstacles. The ego vehicle is equipped with a LiDAR and camera and uses the MEDL-U algorithm [17] as the EDL model. It takes LiDAR point clouds, camera images, and 2D bounding box predictions to predict 3D bounding boxes with uncertainty, represented as NIG distributions. We use YOLOv8 model [21] trained on COCO [22] to predict 2D bounding box and MEDL-U trained on the KITTI dataset [23] using only the car class.

To compare confident and uncertain perception scenarios, we conduct experiments with different obstacle classes and sensor configurations. In the confident case, a car is used as the obstacle, and the LiDAR is mounted on the ego vehicle at the same height as in the training setup, ensuring that the input data remains in-distribution. In contrast, in the uncertain case, a motorcycle is used as the obstacle, and the LiDAR is mounted at a different height, creating out-of-distribution inputs. The EDL model estimates prediction uncertainty, allowing us to analyze how the distributionally robust safety constraint varies with different levels of uncertainty. For simplicity, x and y positions (c_1^ξ, c_2^ξ) are considered with uncertainty distribution and z coordinate is ignored because the height information of ground obstacles has minimal impact on collision avoidance.

We compare the following safety constraint baselines.

- **(Single-Estimate)** uses a single prediction of the obstacle state and thus a deterministic safety constraint.
- **(CVaR [11])** estimates the distribution of the obstacle state $\mathbb{P}_{data}$ using 5 predictions, obtains the distribution of loss $\mathbb{P}_{loss}$ using the unscented transform method, and formulates the CVaR risk metric to the loss distribution.
- **(W-DRO-CVaR [11])** enforces distributionally robust CVaR considering all distributions whose Wasserstein distance from the loss distribution $\mathbb{P}_{loss}$ is within 0.1.
- **(Moment-Based EDL [14])** enforces distributionally robust CVaR, utilizing the variance of the moments $Var[\mu], Var[\sigma^2]$ estimated from the EDL model as a confidence interval of the ambiguity set with a confidence coefficient of 0.5.
- **(DR-EDL-CVaR)** is the proposed distributionally robust safety constraint with $\eta = 0.9$.

All stochastic methods use CVaR with $\epsilon = 0.9$.

We evaluate all methods over 100 runs. Specifically, we measure a collision rate with obstacles and a success rate of reaching the destination without collision. Among successful runs, we also report the average total cost and the average minimum distance to the estimated obstacle center $\|c^x - c^\xi\|_2$, and the average optimization time.

5.2 Results

Table 1: Comparison in confident & uncertain scenarios

Scenario	Method	Succ.(%) ↑	Coll. (%) ↓	Cost ↓	Distance	Time (ms) ↓
Confident	Single-Estimate	100	0	6.576×10^5	2.625	220.5
	CVaR	93	0	6.304×10^{10}	5.528	247.2
	W-DRO-CVaR	29	0	3.786×10^{11}	7.672	1064
	Moment-Based EDL	78	0	2.281×10^7	2.640	219.7
	DR-EDL-CVaR (Ours)	100	0	6.072×10^7	5.032	203.5
Uncertain	Single-Estimate	0	100	-	-	306.4
	CVaR	72	28	1.352×10^{11}	6.756	229.7
	W-DRO-CVaR	65	2	1.144×10^{11}	11.436	1168
	Moment-Based EDL	-	-	-	-	-
	DR-EDL-CVaR (Ours)	95	2	6.831×10^5	8.014	223.1

The results of the confident perception experiments are illustrated in Fig. 3 (a) and Table 1. All methods accurately infer obstacle positions, thereby causing no collisions. The Single-Estimate approach successfully reaches destinations in all runs while maintaining low cost and minimal distance. In contrast, the CVaR and W-DRO-CVaR methods, which consider uncertainty distributions, maintain larger distances from the obstacle. The W-DRO-CVaR method shows a low success rate as it leads to overly conservative behavior and infeasibility due to the use of the unscented transform. The Moment-Based EDL method yields a similar conservative distance to the Single-Estimate approach, but suffers from a low success rate. This is because it is only applicable when perception is highly confident ($\alpha > 2$). Our DR-EDL-CVaR approach formulates a small ambiguity set for confident perception, resulting in less conservative behavior and a higher success rate than CVaR and W-DRO-CVaR.

The results of uncertain perception experiments are shown in Fig. 3 (b) and Table 1. The uncertainty stems from model uncertainty because the perception model encounters a motorbike despite being trained only on the car class. The Single-Estimate and CVaR methods exhibit high collision rates due to their inability to handle model uncertainty. In contrast, W-DRO-CVaR and our proposed method consider model uncertainty using the ambiguity set, resulting in significantly lower collision rates. The Moment-Based EDL method is inapplicable in all uncertain perception scenarios.

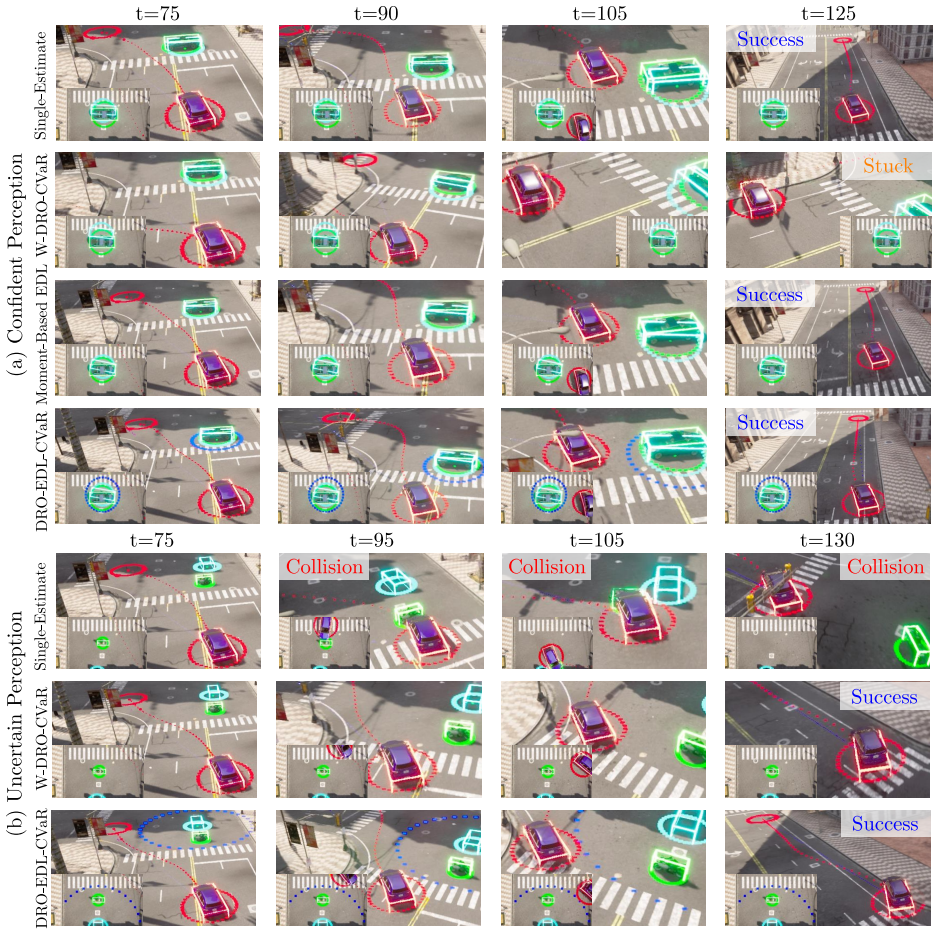

Figure 3: Illustration of the compared methods and our proposed method in (a) confident and (b) uncertain perception scenarios. Ego position (red), obstacle position (green), estimated obstacle position (cyan), and safe boundary (blue) are visualized.

Our DR-EDL-CVaR method is the only one that operates less conservatively under confident perception and more conservatively under uncertain perception, ensuring both efficiency and safety across varying perception confidence levels. These results confirm that our safety constraint formulation proves effective for safe autonomous driving. Additionally, our method demonstrates a 5 times improvement in computational efficiency compared to the W-DRO-CVaR method while comparable to other simpler approaches.

6 Conclusion

We have proposed a safe motion planning approach that integrates EDL-based perception in uncertain environments. Our constraint formulation method, DR-EDL-CVaR, ensures safety through a distributionally robust safety constraint that accounts for the estimated uncertainty. Additionally, we have introduced DRO-EDL-MPC, a conservative yet tractable motion planning algorithm based on the upper bound of the distributional robust safety loss and standardization of the NIG distribution. Experiments demonstrate that our approach is the only method that can produce safe motion in out-of-distribution scenarios without additional training, while exhibiting low conservativeness in in-distribution scenarios and achieving great efficiency compared to other methods. Our future work will address multiple dynamic obstacles.

Acknowledgments

This work was supported by the Hyundai Motor Chung Mong-Koo Foundation.

References

[1] E. Yurtsever, J. Lambert, A. Carballo, and K. Takeda, "A Survey of Autonomous Driving: Common Practices and Emerging Technologies," en, *IEEE Access*, vol. 8, pp. 58 443–58 469, 2020.

[2] C. Sun, R. Zhang, Y. Lu, *et al.*, "Toward Ensuring Safety for Autonomous Driving Perception: Standardization Progress, Research Advances, and Perspectives," en, *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 5, pp. 3286–3304, May 2024.

[3] A. Kendall and Y. Gal, "What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?" en, in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 5580–5590.

[4] A. Nemirovski and A. Shapiro, "Convex Approximations of Chance Constrained Programs," en, *SIAM J. Optim.*, vol. 17, no. 4, pp. 969–996, Jan. 2007.

[5] K. Ren, H. Ahn, and M. Kamgarpour, "Chance-Constrained Trajectory Planning With Multimodal Environmental Uncertainty," en, *IEEE Contr. Syst. Lett.*, vol. 7, pp. 13–18, 2023.

[6] A. Majumdar and M. Pavone, "How Should a Robot Assess Risk? Towards an Axiomatic Theory of Risk in Robotics," en, in *Proc. Int. Symp. Robot. Res.*, Nov. 2017.

[7] A. D. Kiureghian and O. Ditlevsen, "Aleatory or Epistemic? Does It Matter?" en, *Struct. Saf.*, vol. 31, no. 2, pp. 105–112, Mar. 2009.

[8] W. Wiesemann, D. Kuhn, and M. Sim, "Distributionally Robust Convex Optimization," en, *Oper. Res.*, vol. 62, no. 6, pp. 1358–1376, Dec. 2014.

[9] A. Ben-Tal, D. Den Hertog, A. De Waegenaere, B. Melenberg, and G. Rennen, "Robust Solutions of Optimization Problems Affected by Uncertain Probabilities," en, *Manag. Sci.*, vol. 59, no. 2, pp. 341–357, Feb. 2013.

[10] C. Dellaporta, P. O'Hara, and T. Damoulas, "Distributionally Robust Optimisation with Bayesian Ambiguity Sets," en, *arXiv:2409.03492*, 2024.

[11] A. Hakobyan and I. Yang, "Distributionally Robust Optimization with Unscented Transform for Learning-Based Motion Control in Dynamic Environments," en, in *Proc. IEEE Int. Conf. Robot. Automat.*, May 2023, pp. 3225–3232.

[12] S. Safaoui and T. H. Summers, "Distributionally Robust CVaR-Based Safety Filtering for Motion Planning in Uncertain Environments," en, in *Proc. IEEE Int. Conf. Robot. Automat.*, May 2024, pp. 103–109.

[13] E. Delage and Y. Ye, "Distributionally Robust Optimization Under Moment Uncertainty with Application to Data-Driven Problems," en, *Oper. Res.*, vol. 58, no. 3, pp. 595–612, Jun. 2010.

[14] Q. Wang, L. Pan, L. Heistrene, and Y. Levron, "Signal-Devices Management and Data-Driven Evidential Constraints Based Robust Dispatch Strategy of Virtual Power Plant," en, *Expert Syst. Appl.*, vol. 262, p. 125 603, Mar. 2025.

[15] A. Amini, W. Schwarting, A. Soleimany, and D. Rus, "Deep Evidential Regression," en, in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 33, Nov. 2020, pp. 14 927–14 937.

[16] C. Rasmussen, "Gaussian Processes in Machine Learning," en, in *Advanced Lectures on Machine Learning*, Series Title: Lecture Notes in Computer Science, Springer, 2004, pp. 63–71.

[17] H. Paat, Q. Lian, W. Yao, and T. Zhang, "MEDL-U: Uncertainty-Aware 3D Automatic Annotation Based on Evidential Deep Learning," en, in *Proc. IEEE Int. Conf. Robot. Automat.*, Feb. 2024, pp. 13 976–13 982.

[18] J. D. Philip and R. Philip, *Methods of Numerical Integration*, 2nd ed. Courier Corporation, 2007.

[19] P. B. Richard, *Algorithms for Minimization Without Derivatives*. Courier Corporation, 2013.

[20] T. Schouwenaars, B. De Moor, E. Feron, and J. How, "Mixed Integer Programming for Multi-Vehicle Path Planning," en, in *Proc. Eur. Controls Conf.*, Porto, Portugal: IEEE, Sep. 2001, pp. 2603–2608.

[21] J. Glenn, C. Ayush, and Q. Jing, *Ultralytics YOLOv8*, 2023.

[22] T.-Y. Lin, M. Maire, S. Belongie, *et al.*, "Microsoft COCO: Common Objects in Context," en, in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 740–755.

[23] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision Meets Robotics: The KITTI Dataset," en, *Int. J. Robot. Res.*, vol. 32, no. 11, pp. 1231–1237, Sep. 2013.

[24] M. Norton, V. Khokhlov, and S. Uryasev, "Calculating CVaR and bPOE for Common Probability Distributions with Application to Portfolio Optimization and Density Estimation," en, *Ann. Oper. Res.*, vol. 299, no. 1-2, pp. 1281–1315, Apr. 2021.

A Proof of Proposition 1

Proposition 1. (Conservative surrogate ambiguity set solution). *Given the surrogate ambiguity set $\mathbb{D}_i^{sur}(\eta_i|m_i)$, the distributionally robust safety loss satisfies*

$$\max_{\mathbb{P}\in\mathbb{D}(\eta|\mathbf{m})} \mathit{CVaR}_\epsilon^{\mathbb{P}}[\ell(\mathbf{x}(t),\xi)] \leq (r^x+r^\xi)^2 - \sum_i^{n_c} \max_{\mathbb{P}_i\in\mathbb{D}_i^{sur}(\eta_i|m_i)} \mathit{CVaR}_\epsilon^{\mathbb{P}_i}[-|c_i^x - c_i^\xi|]^2.$$

Furthermore,

- *If $c_i^x \in \mathcal{I}_{i,\mu}$,*
$$\max_{\mathbb{P}_i\in\mathbb{D}_i^{sur}(\eta_i|m_i)} \mathit{CVaR}_\epsilon^{\mathbb{P}_i}[-|c_i^x - c_i^\xi|] = \kappa\cdot\sigma_{i,min} < 0,$$
where $\kappa = \frac{1}{1-\epsilon}\sqrt{\frac{2}{\pi}}(\exp(-[erf^{-1}(\epsilon-1)]^2)-1)$ and $erf^{-1}(\cdot)$ the inverse of the error function.

- *If $c_i^x \in \mathcal{I}_{i,unsafe}\setminus\mathcal{I}_{i,\mu}$,*
$$\max_{\mathbb{P}_i\in\mathbb{D}_i^{sur}(\eta_i|m_i)} \mathit{CVaR}_\epsilon^{\mathbb{P}_i}[-|c_i^x - c_i^\xi|] \leq \kappa\cdot\sigma_{i,min} < 0,$$
where $\mathcal{I}_{i,unsafe} = [\mu_{i,min} - \delta\cdot\sigma_{i,max}, \mu_{i,max} + \delta\cdot\sigma_{i,max}]$ for $\delta = \frac{1}{1-\epsilon}\frac{1}{\sqrt{2\pi}}\exp(-[erf^{-1}(2\epsilon-1)]^2) > 0$.

- *If $c_i^x \notin \mathcal{I}_{i,unsafe}$,*
$$\max_{\mathbb{P}_i\in\mathbb{D}_i^{sur}(\eta_i|m_i)} \mathit{CVaR}_\epsilon^{\mathbb{P}_i}[-|c_i^x - c_i^\xi|] < -|c_i^x-\gamma_i| + \frac{\mu_{i,max}-\mu_{i,min}}{2} + \delta\cdot\sigma_{i,max} < 0,$$
where γ_i is given in the prediction result m_i.

Proof. The worst-case CVaR of the safety loss (2) is

$$\begin{aligned}
&\max_{\mathbb{P}\in\mathbb{D}(\eta|\mathbf{m})} \mathrm{CVaR}_\epsilon^{\mathbb{P}}[\ell(\mathbf{x}(t),\xi)] \\
&= (r^x+r^\xi)^2 + \max_{\mathbb{P}\in\mathbb{D}(\eta|\mathbf{m})} \mathrm{CVaR}_\epsilon^{\mathbb{P}}[-\sum_i^{n_c}|c_i^x - c_i^\xi|^2] \\
&\leq (r^x+r^\xi)^2 - \sum_i^{n_c} \max_{\mathbb{P}_i\in\mathbb{D}_i^{sur}(\eta_i|m_i)} \mathrm{CVaR}_\epsilon^{\mathbb{P}_i}[-|c_i^x - c_i^\xi|]^2
\end{aligned} \tag{11}$$

The equality holds as the axes of the obstacle distribution are independent of each other and by Lemma 2(a) and (8).

We will derive the upper bound of $\max_{\mathbb{P}_i\in\mathbb{D}_i^{sur}(\eta_i|m_i)} \mathrm{CVaR}_\epsilon^{\mathbb{P}_i}[-|c_i^x - c_i^\xi|]$ that is strictly negative and thus minimizes the squared term in (11). The difference $X_d = c_i^x - c_i^\xi$ follows the normal distribution

$\mathcal{N}(c_i^x - \mu_i, \sigma_i^2)$ when c_i^ξ follows $\mathbb{P}_i = \mathcal{N}(\mu_i, \sigma_i^2)$. We denote $\mu_d := c_i^x - \mu_i$ and $\sigma_d^2 := \sigma_i^2$. The distance $|X_d|$ follows a folded normal distribution. By Lemma 2(b), $\text{CVaR}_\epsilon^{\mathbb{P}_i}[-|X_d|]$, referred to as the *negative distance CVaR*, is a monotonically decreasing function with respect to $|\mu_d|$, and thus, its maximum is attained when $|\mu_d|$ is minimized.

If $c_i^x \in \mathcal{I}_{i,\mu}$, $|\mu_d| = 0$ maximizes the negative distance CVaR. The distribution of $-|X_d|$ is a half-normal distribution, and the negative distance CVaR is $\kappa \cdot \sigma_i$ by Lemma 2(c). Therefore, the distribution $\mathbb{P}_i = \mathcal{N}(c_i^x, \sigma_{i,min}^2)$, which lies in the surrogate ambiguity set $\mathbb{D}_i^{sur}(\eta_i|m_i)$, yields the maximum negative distance CVaR of $\kappa \cdot \sigma_{i,min}$.

If $c_i^x \in \mathcal{I}_{i,unsafe} \backslash \mathcal{I}_{i,\mu}$, the mean $|\mu_d|$ of $|X_d|$ is larger than 0 because the mean μ_i of c_i^ξ is confined in $\mathcal{I}_{i,\mu}$. Because the maximum negative distance CVaR decreases with respect to $|\mu_d|$ by Lemma 2(b), it is no larger than $\kappa \cdot \sigma_{i,min}$.

If $c_i^x \notin \mathcal{I}_{i,unsafe}$, when $\mu_d > 0$, the ϵ-quantile of the distribution of X_d, i.e., $\text{VaR}_\epsilon^{\mathbb{P}_i}[X_d]$, is a positive value by construction of $\mathcal{I}_{i,unsafe}$, in particular the definition of δ. Similary, when $\mu_d < 0$, $\text{VaR}_\epsilon^{\mathbb{P}_i}[X_d] < 0$ by construction of $\mathcal{I}_{i,unsafe}$. Under this condition, Lemma 2(d) tells that $\max_{\mathbb{P}_i} -|\mu_d| + \delta \cdot \sigma_d > \max_{\mathbb{P}_i} \text{CVaR}_\epsilon^{\mathbb{P}_i}[-|X_d|]$. The maximum of $-|\mu_d| + \delta \cdot \sigma_d$ is attained at $\sigma_d = \sigma_{i,max}$ because $\delta > 0$ and $|\mu_d| = |c_i^x - \mu_{i,min}|$ when $c_i^x < \mu_{i,min} - \delta\sigma_{i,max}$ or at $\sigma_d = \sigma_{i,max}$ and $|\mu_d| = |c_i^x - \mu_{i,max}|$ when $c_i^x > \mu_{i,\max} + \delta \cdot \sigma_{i,max}$. Using γ_i, which is the center of $[\mu_{i,min}, \mu_{i,max}]$, we can represent such optimal $|\mu_d|$ as $|c_i^x - \gamma_i| - (\mu_{i,max} - \mu_{i,min})/2$. Therefore, $-|c_i^x - \gamma_i| + (\mu_{i,max} - \mu_{i,min})/2 + \delta \cdot \sigma_{i,max} > \max_{\mathbb{P}_i} \text{CVaR}_\epsilon^{\mathbb{P}_i}[-|X_d|]$ and this upper bound is negative because $c_i^x \notin \mathcal{I}_{i,unsafe}$ and by construction of $\mathcal{I}_{i,unsafe}$. □

B Proof of Proposition 2

Proposition 2. (Constraint reformulation). *Given the perception results* $\mathbf{m}, \phi^\xi, v^\xi$, *and* $\mathbf{b}^\xi$, *define* $\xi_{EDL} = (\mathbf{c}^{\xi_{EDL}}, \phi^{\xi_{EDL}}, v^{\xi_{EDL}})$ *as*

$$c_i^{\xi_{EDL}} := \gamma_i, \;\; \phi^{\xi_{EDL}} := \phi^\xi, \;\; v^{\xi_{EDL}} := v^\xi,$$

and its attribute $\mathbf{b}^{\xi_{EDL}}$ *as*

$$b_i^{\xi_{EDL}} := (\mu_{i,max} - \mu_{i,min})/2 + \delta \cdot \sigma_{i,max} + r^\xi,$$

where δ *is provided in Proposition 1. Then, for any* $\mathbf{x}$,

$$\ell(\mathbf{x}, \xi_{EDL}) \leq 0 \implies \textit{DR-EDL-CVaR}_\epsilon^{\mathbb{D}(\eta|\mathbf{m})}[\ell(\mathbf{x}, \xi)] \leq 0. \tag{12}$$

Proof. Suppose $c_i^x \in \mathcal{I}_{i,unsafe}$ for $i = 1, \ldots, n_{in}$ and $c_i^x \notin \mathcal{I}_{i,unsafe}$ for $i \in n_{in} + 1, \ldots, n_c$. By Proposition 1,

$$\begin{aligned} &\max_{\mathbb{P} \in \mathbb{D}(\eta|\mathbf{m})} \text{CVaR}_\epsilon^{\mathbb{P}}[\ell(\mathbf{x}, \xi)] \\ \leq &(r^x + r^\xi)^2 - \sum_{i=1}^{n_{in}} (\kappa \cdot \sigma_{i,min})^2 - \sum_{i=n_{in}+1}^{n_c} (|c_i^x - \gamma_i| - h_i)^2, \end{aligned}$$

where $h_i = (\mu_{i,max} - \mu_{i,min})/2 + \delta \cdot \sigma_{i,max}$. We denote the right-hand side by $f(\mathbf{x})$.

Let us define a new obstacle state $\tilde{\xi} = (c^{\tilde{\xi}}, \phi^{\tilde{\xi}}, v^{\tilde{\xi}})$ with $c^{\tilde{\xi}} = \gamma_i, \phi^{\tilde{\xi}} = \phi^\xi, v^{\tilde{\xi}} = v^\xi, ||b_i^{\tilde{\xi}}||_2^2 = r^\xi + \sqrt{\sum_i^{n_c} h_i^2}$, and let

$$g(\mathbf{x}) := \ell(\mathbf{x}, \tilde{\xi}) = (r^x + r^\xi + \sqrt{\textstyle\sum_i^{n_c} h_i^2})^2 - \textstyle\sum_i^{n_c} (|c_i^x - \gamma_i|)^2.$$

Note that $\tilde{\xi}$ is the same as ξ_{EDL} except that the attribute $b_i^{\xi_{EDL}} = r^\xi + h_i$ incorporates the uncertainty margin h_i directly into the size of the obstacle.

Given feasible sets $F := \{\mathbf{x} | f(\mathbf{x}) \leq 0\}$ and $G := \{\mathbf{x} | g(\mathbf{x}) \leq 0\}$, we aim to show $G \subseteq F$. Let $\vec{v}(\mathbf{x}) := (|c_1^x - \gamma_1|, \ldots, |c_{n_c}^x - \gamma_{n_c}|), \vec{h} := (h_1, \ldots, h_{n_c})$ and $s(\mathbf{x}) := (\min(h_1, |c_1^x -$

$\gamma_1|), ..., \min(h_{n_c}, |c^x_{n_c} - \gamma_{n_c}|))$. Then,

$$\begin{aligned}\|\vec{v}(\mathbf{x})\|_2 &\leq \|\vec{s}(\mathbf{x})\|_2 + \|\vec{v}(\mathbf{x}) - \vec{s}(\mathbf{x})\|_2 \\ &\leq \|\vec{h}\|_2 + \sqrt{\textstyle\sum_{i=n_{in}+1}^{n_c} (|c^x_i - \gamma_i| - h_i)^2}. \end{aligned} \tag{13}$$

The first inequality is by the triangle inequality, and the second inequality is because $\|\vec{s}(\mathbf{x})\|_2 \leq \|\vec{h}\|_2$ by definition and $s_i(\mathbf{x}) = |c^x_i - \gamma_i|$ for $i = 1, \ldots, n_{in}$ and h_i for $i = n_{in} + 1, \ldots, n_c$.

Now, if $f(\mathbf{x}) > 0$, we have $(r^x + r^\xi)^2 - \sum_{i=n_{in}+1}^{n_c}(|c^x_i - \gamma_i| - h_i)^2 > 0$ because $(\kappa \cdot \sigma_{i,min})^2 \geq 0$. Therefore, (13) becomes $\|\vec{v}(\mathbf{x})\|_2 < \|\vec{h}\|_2 + (r^x + r^\xi)$, and this leads to $g(\mathbf{x}) > 0$. This showed $F^c \subseteq G^c$ and therefore $G \subseteq F$. This implies that $g(\mathbf{x}) \leq 0$ is more conservative than the original constraint $f(\mathbf{x}) \leq 0$.

To facilitate a simpler and more tractable representation of the obstacle attribute, we define the obstacle ξ_{EDL} with $b_i^{\xi_{EDL}} = r^\xi + h_i$. This makes the radius $r^{\xi_{EDL}} = \|r^\xi + \vec{h}\|_2 \geq r^\xi + \|\vec{h}\|_2$ and thus $\ell(\mathbf{x}, \xi_{EDL}) \geq \ell(\mathbf{x}, \tilde{\xi})$. Therefore, $\ell(\mathbf{x}, \xi_{EDL}) \leq 0 \implies \ell(\mathbf{x}, \tilde{\xi}) \leq 0 \implies f(\mathbf{x}) \leq 0 \implies$ DR-EDL-CVaR$_\epsilon^{\mathbb{D}(\eta|\mathbf{m})}[\ell(\mathbf{x}, \xi)] \leq 0$. □

C Properties of the Value-at-Risk

Lemma 1. (Properties of the Value-at-Risk). *Given a random variable $X \sim \mathbb{P} = \mathcal{N}(\mu, \sigma^2)$ and a threshold $\epsilon \in [0, 1)$, the following holds.*

(a) $\mathit{VaR}_\epsilon^{\mathbb{P}}[-|X|^2] = -\mathit{VaR}_\epsilon^{\mathbb{P}}[-|X|]^2$

(b) For $Y = -|X|$ with a fixed σ^2,

$$\frac{d\mathit{VaR}_\epsilon^{\mathbb{P}}[Y]}{d\mu} = \frac{\phi(\frac{\mathit{VaR}_\epsilon^{\mathbb{P}}[Y]-\mu}{\sigma}) - \phi(\frac{-\mathit{VaR}_\epsilon^{\mathbb{P}}[Y]-\mu}{\sigma})}{\phi(\frac{\mathit{VaR}_\epsilon^{\mathbb{P}}[Y]-\mu}{\sigma}) + \phi(\frac{-\mathit{VaR}_\epsilon^{\mathbb{P}}[Y]-\mu}{\sigma})},$$

where ϕ is the probability density function of the standard normal distribution.

(c) $\mathit{VaR}_\epsilon^{\mathbb{P}}[-|X|] = \sqrt{2}\sigma \cdot erf^{-1}(\epsilon - 1)$

(d) $\mathit{VaR}_\epsilon^{\mathbb{P}}[X] > \mathit{VaR}_\epsilon^{\mathbb{P}}[-|X|]$

Proof. (a)

$$\begin{aligned} &P[-|X| \leq \text{VaR}_\epsilon^{\mathbb{P}}[-|X|]] = \epsilon, \\ &P[-|X|^2 \leq \text{VaR}_\epsilon^{\mathbb{P}}[-|X|^2]] = \epsilon \\ \Leftrightarrow &P[|X| \geq \sqrt{-\text{VaR}_\epsilon^{\mathbb{P}}[-|X|^2]}] = \epsilon \\ \Leftrightarrow &P[-|X| \leq -\sqrt{-\text{VaR}_\epsilon^{\mathbb{P}}[-|X|^2]}] = \epsilon. \end{aligned}$$

$$\therefore \text{VaR}_\epsilon^{\mathbb{P}}[-|X|^2] = -\text{VaR}_\epsilon^{\mathbb{P}}[-|X|]^2$$

(b) Let $F_Y(y)$ and $f_Y(y)$ be the cumulative probability distribution and the probability density function of a random variable Y, respectively. Let us similarly define $F_X(x)$ and $f_X(x)$ for a random variable X. Also, let Φ be the cumulative distribution function of the standard normal distribution. For simplicity, we denote $\text{Var}_\epsilon^{\mathbb{P}}[Y]$ by $k(\mu)$, which is also a function of σ but we consider it as a constant.

By definition of VaR, we have

$$\begin{aligned}\epsilon = F_Y(k(\mu)) &= \int_{-\infty}^{k(\mu)} f_Y(y)dy \\ &= \int_{-\infty}^{k(\mu)} (f_X(x) + f_X(-x))dx \\ &= \int_{-\infty}^{k(\mu)} f_X(x)dx + \int_{-\infty}^{k(\mu)} f_X(-x)dx.\end{aligned}$$

Let $t = -x$. Then,

$$\begin{aligned}F_Y(k(\mu)) &= \int_{-\infty}^{k(\mu)} f_X(x)dx + \int_{\infty}^{-k(\mu)} f_X(t)(-dt) \\ &= \int_{-\infty}^{k(\mu)} f_X(x)dx - \int_{\infty}^{-k(\mu)} f_X(x)dx \\ &= \Phi\left(\frac{k(\mu)-\mu}{\sigma}\right) - \Phi\left(\frac{-k(\mu)-\mu}{\sigma}\right) + 1.\end{aligned}$$

Because $\epsilon = F_Y(k(\mu))$, we have $\frac{dF_Y(k(\mu))}{d\mu} = 0$. This leads to

$$\phi\left(\frac{k(\mu)-\mu}{\sigma}\right) \cdot \left(\frac{dk(\mu)}{d\mu} - 1\right)\frac{1}{\sigma} - \phi\left(\frac{-k(\mu)-\mu}{\sigma}\right) \cdot \left(-\frac{dk(\mu)}{d\mu} - 1\right)\frac{1}{\sigma} = 0.$$

Then

$$\frac{dk(\mu)}{d\mu} = \frac{\phi\left(\frac{k(\mu)-\mu}{\sigma}\right) - \phi\left(\frac{-k(\mu)-\mu}{\sigma}\right)}{\phi\left(\frac{k(\mu)-\mu}{\sigma}\right) + \phi\left(\frac{-k(\mu)-\mu}{\sigma}\right)}.$$

This concludes the proof.

(c) Let $-|X| = Y$. Then the probability density function of Y is

$$f_Y(y) = f_X(x) + f_X(-x) = \frac{\sqrt{2}}{\sigma\sqrt{\pi}} exp(-\frac{y^2}{2\sigma^2}), y \leq 0.$$

And the cumulative distribution function of Y is

$$F_Y(y) = \int_{-\infty}^{y} f_Y(t)dt = \int_{-\infty}^{y} \frac{\sqrt{2}}{\sigma\sqrt{\pi}} exp(-\frac{t^2}{2\sigma^2})dt.$$

Meanwhile,

$$F_Y(0) = \int_{-\infty}^{0} f_Y(t)dt = 1$$

because Y is defined in $(-\infty, 0]$. Then

$$\begin{aligned}F_Y(0) &= \int_{-\infty}^{y} f_Y(t)dt + \int_{y}^{0} f_Y(t)dt \\ &= \int_{-\infty}^{y} f_Y(t)dt - \int_{0}^{y} f_Y(t)dt.\end{aligned}$$

Therefore,

$$\begin{aligned}\int_{-\infty}^{y} f_Y(t)dt &= 1 + \int_{0}^{y} f_Y(t)dt \\ &= 1 + \int_{0}^{y} \frac{\sqrt{2}}{\sigma\sqrt{\pi}} exp(-\frac{t^2}{2\sigma^2})dt.\end{aligned}$$

Let $t/\sqrt{2}\sigma = k$. Then

$$\begin{aligned} F_Y(y) &= 1 + \frac{\sqrt{2}}{\sigma\sqrt{\pi}} \int_0^{y/\sqrt{2}\sigma} exp(-k^2) \cdot \sqrt{2}\sigma dk \\ &= 1 + \frac{2}{\sqrt{\pi}} \int_0^{y/\sqrt{2}\sigma} exp(-k^2) dk \\ &= 1 + erf(y/\sqrt{2}\sigma). \end{aligned}$$

$\text{VaR}_\epsilon^{\mathbb{P}}[Y]$ is the value of y when $F_Y(y) = \epsilon$. Therefore,

$$\text{VaR}_\epsilon^{\mathbb{P}}[Y] = \sqrt{2}\sigma \cdot erf^{-1}(\epsilon - 1).$$

(d) Let $\text{VaR}_\epsilon^{\mathbb{P}}[X] = k, \text{VaR}_\epsilon^{\mathbb{P}}[-|X|] = k'$. By definition,

$$\int_k^\infty f_X(x)dx = 1 - \epsilon = \int_{k'}^0 f_{-|X|}(x)dx.$$

Also,

$$\begin{aligned} \int_k^\infty f_X(x)dx &= \int_k^0 f_X(x)dx + \int_0^{-k} f_X(x)dx + \int_{-k}^\infty f_X(x)dx \\ &= \int_k^0 f_X(x)dx + \int_k^0 f_X(-x)dx + \int_{-k}^\infty f_X(x)dx \\ &= \int_k^0 f_{-|X|}(x)dx + \int_{-k}^\infty f_X(x)dx. \end{aligned}$$

Therefore,

$$\int_{-k}^\infty f_X(x)dx > 0 \implies \int_{k'}^0 f_{-|X|}(x)dx - \int_k^0 f_{-|X|}(x)dx = \int_{k'}^k f_{-|X|}(x)dx > 0.$$

Therefore, $\text{VaR}_\epsilon^{\mathbb{P}}[X] > \text{VaR}_\epsilon^{\mathbb{P}}[-|X|]$. □

D Properties of the Conditional Value-at-Risk

Lemma 2. (Properties of the Conditional Value-at-Risk). *Given a random variable $X \sim \mathbb{P} = \mathcal{N}(\mu, \sigma^2)$ and a threshold $\epsilon \in [0, 1)$, the following holds.*

(a) $CVaR_\epsilon^{\mathbb{P}}[-|X|^2] \leq -CVaR_\epsilon^{\mathbb{P}}[-|X|]^2$

(b) For a fixed σ^2, $CVaR_\epsilon^{\mathbb{P}}[-|X|]$ is monotonically decreasing with respect to $|\mu|$.

(c) If $\mu = 0$, then $CVaR_\epsilon^{\mathbb{P}}[-|X|] = \kappa \cdot \sigma$,
where $\kappa = \frac{1}{1-\epsilon}\sqrt{2/\pi}[exp(-[erf^{-1}(\epsilon - 1)]^2) - 1]$.

(d) If $VaR_\epsilon^{\mathbb{P}}[X] > 0 \wedge \mu > 0$, or $VaR_\epsilon^{\mathbb{P}}[X] < 0 \wedge \mu < 0$, then $-|\mu| + \delta \cdot \sigma > CVaR_\epsilon^{\mathbb{P}}[-|X|]$.

Proof. (a)

$$\begin{aligned} \text{CVaR}_\epsilon^{\mathbb{P}}[-|X|^2] &= \mathbb{E}[-|X|^2 : -|X|^2 \geq \text{VaR}_\epsilon^{\mathbb{P}}[-|X|^2]] && (14) \\ &= \mathbb{E}[-|X|^2 : -|X|^2 \geq -\text{VaR}_\epsilon^{\mathbb{P}}[-|X|]^2] && (15) \\ &= \mathbb{E}[-|X|^2 : |X| \leq -\text{VaR}_\epsilon^{\mathbb{P}}[-|X|]] && (16) \\ &= -\mathbb{E}[|X|^2 : |X| \leq -\text{VaR}_\epsilon^{\mathbb{P}}[-|X|]] && (17) \\ &\leq -\mathbb{E}[-|X| : |X| \leq -\text{VaR}_\epsilon^{\mathbb{P}}[-|X|]]^2 && (18) \\ &= -\mathbb{E}[-|X| : -|X| \geq \text{VaR}_\epsilon^{\mathbb{P}}[-|X|]]^2 && (19) \\ &= -\text{CVaR}_\epsilon^{\mathbb{P}}[-|X|]^2. && (20) \end{aligned}$$

Equation (15) holds by Lemma 1 (a), and inequality (18) holds by Jensen's inequality.

(b) With the same notations used in the proof of Lemma 1 (b), we have

$$\begin{aligned}\text{CVaR}_\epsilon^{\mathbb{P}}[Y] = \mathbb{E}[Y : Y \geq k(\mu)] &= \frac{1}{1-\epsilon}\int_{k(\mu)}^{0} y f_Y(y)dy \\ &= \frac{1}{1-\epsilon}\int_{k(\mu)}^{0} x(f_X(x)+f_X(-x))dx \\ &= \frac{1}{1-\epsilon}\left[\int_{k(\mu)}^{0} x f_X(x)dx + \int_{-k(\mu)}^{0} t f_X(t)dt\right] \\ &= \frac{1}{1-\epsilon}\left[\int_{k(\mu)}^{0} x f_X(x)dx - \int_{0}^{-k(\mu)} x f_X(x)dx\right].\end{aligned}$$

Let $a(\mu), b(\mu)$ be functions with respect to μ, and standardized random variable $Z = (X-\mu)/\sigma$. Then,

$$\int_{a(\mu)}^{b(\mu)} x f_X(x)dx = \int_{\frac{a(\mu)-\mu}{\sigma}}^{\frac{b(\mu)-\mu}{\sigma}} (\mu+\sigma\cdot z)\frac{1}{\sigma}\phi(z)\sigma dz = \mu\int_{\frac{a(\mu)-\mu}{\sigma}}^{\frac{b(\mu)-\mu}{\sigma}} \phi(z)dz + \sigma\int_{\frac{a(\mu)-\mu}{\sigma}}^{\frac{b(\mu)-\mu}{\sigma}} z\phi(z)dz.$$

With $z_a(\mu) = \frac{a(\mu)-\mu}{\sigma}$ and $z_b(\mu) = \frac{b(\mu)-\mu}{\sigma}$, the derivative with respect to μ is

$$\begin{aligned}\frac{d}{d\mu}\int_{a(\mu)}^{b(\mu)} x f_X(x)dx &= \left[\int_{z_a(\mu)}^{z_b(\mu)} \phi(z)dz + \mu\cdot\{\phi(z_b(\mu))\cdot z_b'(\mu) - \phi(z_a(\mu))\cdot z_a'(\mu)\}\right] \\ &\quad + \sigma\left[z_b(\mu)\phi(z_b(\mu))z_b'(\mu) - z_a(\mu)\phi(z_a(\mu))z_a'(\mu)\right] \\ &= \int_{z_a(\mu)}^{z_b(\mu)} \phi(z)dz + b(\mu)\phi(z_b(\mu))z_b'(\mu) - a(\mu)\phi(z_a(\mu))z_a'(\mu).\end{aligned}$$

Using this result and letting $z_k(\mu) = \frac{k(\mu)-\mu}{\sigma}, z_{-k}(\mu) = \frac{-k(\mu)-\mu}{\sigma}$ and $z_0(\mu) = \frac{-\mu}{\sigma}$, we can express the derivative of $\text{CVaR}_\epsilon^{\mathbb{P}}[Y]$ with respect to μ as follows.

$$\begin{aligned}(1-\epsilon)\frac{d}{d\mu}\text{CVaR}_\epsilon^{\mathbb{P}}[Y] &= \frac{d}{d\mu}\left[\int_{k(\mu)}^{0} x f_X(x)dx - \int_{0}^{-k(\mu)} x f_X(x)dx\right] \\ &= \int_{z_k(\mu)}^{z_0(\mu)} \phi(z)dz - k(\mu)\cdot\phi(z_k(\mu))\cdot z_k'(\mu) - \int_{z_0(\mu)}^{z_{-k}(\mu)} \phi(z)dz - \left[(-k(\mu))\cdot\phi(z_{-k}(\mu))\cdot z_{-k}'(\mu)\right] \\ &= \int_{z_k(\mu)}^{z_0(\mu)} \phi(z)dz - \int_{z_0(\mu)}^{z_{-k}(\mu)} \phi(z)dz - k(\mu)\left[\phi(z_k(\mu))\cdot z_k'(\mu) - \phi(z_{-k}(\mu))\cdot z_{-k}'(\mu)\right].\end{aligned}$$

Because

$$z_k'(\mu) = \frac{k'(\mu)-1}{\sigma}, z_{-k}'(\mu) = \frac{-k'(\mu)-1}{\sigma},$$

and by Lemma 1 (b),

$$k'(\mu) = \frac{dk(\mu)}{d\mu} = \frac{\phi(z_k(\mu)) - \phi(z_{-k}(\mu))}{\phi(z_k(\mu)) + \phi(z_{-k}(\mu))},$$

we have $\phi(z_k(\mu))\cdot z_k'(\mu) - \phi(z_{-k}(\mu))\cdot z_{-k}'(\mu) = 0$. Therefore,

$$\frac{d}{d\mu}\text{CVaR}_\epsilon^{\mathbb{P}}[Y] = \frac{1}{1-\epsilon}\left[\int_{\frac{k(\mu)-\mu}{\sigma}}^{\frac{-\mu}{\sigma}} \phi(z)dz - \int_{\frac{-\mu}{\sigma}}^{\frac{-k(\mu)-\mu}{\sigma}} \phi(z)dz\right].$$

Note that for any $b \leq 0$, $\int_{a+b}^{a}\phi(z)dz > \int_{a}^{a-b}\phi(z)dz$ if $a > 0$ and $\int_{a+b}^{a}\phi(z)dz < \int_{a}^{a-b}\phi(z)dz$ if $a < 0$. Therefore,

$$\frac{d}{d\mu}\text{CVaR}_\epsilon^{\mathbb{P}}[Y] < 0 \text{ for } \mu > 0.$$

and

$$\frac{d}{d\mu}\text{CVaR}_\epsilon^{\mathbb{P}}[Y] > 0 \text{ for } \mu < 0.$$

That is, $\text{CVaR}_\epsilon^{\mathbb{P}}[Y]$ is monotonically decreasing with respect to $|\mu|$.

(c) Let $-|X| = Y$ and $\text{VaR}_\epsilon[Y] = k$. Then

$$\begin{aligned}
\text{CVaR}_\epsilon^{\mathbb{P}}[Y] &= \mathbb{E}[Y : Y \geq k] \\
&= \frac{1}{1-\epsilon}\int_k^0 y f_Y(y)dy \\
&= \frac{1}{1-\epsilon}\int_k^0 y \cdot \frac{\sqrt{2}}{\sqrt{\pi}\sigma} exp(-\frac{y^2}{2\sigma^2})dy.
\end{aligned}$$

Let $y/\sqrt{2}\sigma = u, y = \sqrt{2}\sigma u, dy = \sqrt{2}\sigma du$. Then

$$\begin{aligned}
&= \frac{1}{1-\epsilon}\int_{k/\sqrt{2}\sigma}^0 \sqrt{2}\sigma u \cdot \frac{\sqrt{2}}{\sqrt{\pi}\sigma} exp(-u^2)\sqrt{2}\sigma du \\
&= \frac{1}{1-\epsilon} \cdot \frac{2\sqrt{2}}{\sqrt{\pi}}\sigma \int_{k/\sqrt{2}\sigma}^0 u \cdot exp(-u^2)du \\
&= \frac{1}{1-\epsilon} \cdot \frac{2\sqrt{2}}{\sqrt{\pi}}\sigma \left[-\frac{1}{2}exp(-u^2)\right]_{k/\sqrt{2}\sigma}^0 \\
&= \frac{1}{1-\epsilon} \cdot \frac{\sqrt{2}}{\sqrt{\pi}}\sigma \left[exp(-\left[k/\sqrt{2}\sigma\right]^2) - 1\right] \\
&= \frac{1}{1-\epsilon} \cdot \frac{\sqrt{2}}{\sqrt{\pi}}\sigma \left[exp(-\left[\frac{\sqrt{2}\sigma \cdot erf^{-1}(\epsilon-1)}{\sqrt{2}\sigma}\right]^2) - 1\right] \\
&= \frac{1}{1-\epsilon} \cdot \frac{\sqrt{2}}{\sqrt{\pi}}\sigma \left[exp(-[erf^{-1}(\epsilon-1)]^2) - 1\right] \\
&= \kappa \cdot \sigma,
\end{aligned}$$

where $k = \sqrt{2}\sigma \cdot erf^{-1}(\epsilon - 1)$ by Lemma 1 (c) and $\kappa = \frac{1}{1-\epsilon}\sqrt{2/\pi}\left[exp(-[erf^{-1}(\epsilon-1)]^2) - 1\right]$.

(d) If $\mu < 0$ and $\text{VaR}_\epsilon^{\mathbb{P}}[X] = k < 0$, then

$$
\begin{aligned}
&\text{CVaR}_\epsilon^{\mathbb{P}}[X] = \mathbb{E}[X|X \geq k] \\
&= \frac{1}{1-\epsilon}\int_k^{\infty} x f_X(x)dx \\
&> \frac{1}{1-\epsilon}\int_k^{-k} x f_X(x)dx \\
&= \frac{1}{1-\epsilon}[\int_k^{0} x f_X(x)dx + \int_0^{-k} x f_X(x)dx] \\
&= \frac{1}{1-\epsilon}[\int_k^{0} x f_X(x)dx + \int_0^{k} x f_X(-x)dx] \\
&= \frac{1}{1-\epsilon}[\int_k^{0} x f_X(x)dt - \int_k^{0} x f_X(-x)dx] \\
&\geq \frac{1}{1-\epsilon}[\int_k^{0} x f_X(x)dt + \int_k^{0} x f_X(-x)dx] \\
&= \frac{1}{1-\epsilon}\int_k^{0} x(f_X(-x) + f_X(x))dx \\
&= \frac{1}{1-\epsilon}\int_k^{0} x f_{-|X|}(x)dx \\
&\geq \frac{1}{1-\epsilon}\int_{\text{VaR}_\epsilon^{\mathbb{P}}[-|X|]}^{0} x f_{-|X|}(x)dx \\
&= \mathbb{E}[-|X| : -|X| \geq \text{VaR}_\epsilon^{\mathbb{P}}[-|X|]] \\
&= \text{CVaR}_\epsilon^{\mathbb{P}}[-|X|].
\end{aligned}
$$

The last inequality is because $k > \text{VaR}_\epsilon^{\mathbb{P}}[-|X|]$ by Lemma 1 (d). Therefore, $\text{CVaR}_\epsilon^{\mathbb{P}}[X] = \mu + \delta \cdot \sigma > \text{CVaR}_\epsilon^{\mathbb{P}}[-|X|]$.

Similarly, if $\mu > 0$ and $\text{VaR}_\epsilon^{\mathbb{P}}[X] > 0$, the mean of $-X$ is smaller than zero and $\text{VaR}_\epsilon^{\mathbb{P}}[-X] < 0$. Thus, we can use the inequality that we derived in the preceding paragraph, that is,

$$\text{CVaR}_\epsilon^{\mathbb{P}}[-X] > \text{CVaR}_\epsilon^{\mathbb{P}}[-|X|].$$

Because X has a symmetric distribution about the mean μ, we have $-\text{VaR}_\epsilon^{\mathbb{P}}[-X] = \text{VaR}_{1-\epsilon}^{\mathbb{P}}[X]$. Also, $\text{CVaR}_{1-\epsilon}^{\mathbb{P}}[X] = \mathbb{E}[X : X \geq \text{VaR}_\epsilon^{\mathbb{P}}[X]] = \mu - \delta \cdot \sigma$ by [24]. Thus,

$$
\begin{aligned}
\text{CVaR}_\epsilon^{\mathbb{P}}[-X] &= \mathbb{E}[-X : -X \geq \text{VaR}_\epsilon^{\mathbb{P}}[-X]] \\
&= -\mathbb{E}[X : X \leq -\text{VaR}_\epsilon^{\mathbb{P}}[-X]] \\
&= -\mathbb{E}[X : X \leq \text{VaR}_{1-\epsilon}^{\mathbb{P}}[X]] \\
&= -(\mu - \delta \cdot \sigma).
\end{aligned}
$$

Therefore, if $\text{VaR}_\epsilon^{\mathbb{P}}[X] > 0 \wedge \mu > 0$ or $\text{VaR}_\epsilon^{\mathbb{P}}[X] < 0 \wedge \mu < 0$, then $-|\mu| + \delta \cdot \sigma > \text{CVaR}_\epsilon^{\mathbb{P}}[-|X|]$. $\square$