Here is a tweaked formulation compared to the InferOpt paper, which should make everything simpler.

Consider a linear solver

$$f(\theta) = \arg\max_{v \in \mathcal{V}} \theta^\top v$$

We define the multiplicative perturbation, where every expectation is taken over $Z \sim \mathcal{N}(0, I)$:

$$\widehat{f}_\varepsilon^\odot(\theta) = \mathbb{E}\left[\arg\max_{v \in \mathcal{V}} \left(\theta \odot e^{\varepsilon Z - \varepsilon^2/2}\right)^\top v\right] = \mathbb{E}\left[f\left(\theta \odot e^{\varepsilon Z - \varepsilon^2/2}\right)\right]$$

The scaling parameter is chosen that way because $\mathbb{E}\left[e^{\varepsilon Z - \varepsilon^2/2}\right] = 1$.

To exhibit a Fenchel-Young loss, we choose the regularization as $\Omega_\varepsilon^\odot = (F_\varepsilon^\odot)^*$ where

$$F_\varepsilon^\odot(\theta) = \mathbb{E}\left[e^{-(\varepsilon Z - \varepsilon^2/2)} \odot \max_{v \in \mathcal{V}} \left(\theta \odot e^{\varepsilon Z - \varepsilon^2/2}\right)^\top v\right]$$

We define

$$f_{\varepsilon,Z}(\theta, v) = \left(\theta \odot e^{\varepsilon Z - \varepsilon^2/2}\right)^\top v$$

so that

$$F_\varepsilon^\odot(\theta) = \mathbb{E}\left[e^{-(\varepsilon Z - \varepsilon^2/2)} \odot \max_{v \in \mathcal{V}} f_{\varepsilon,Z}(\theta, v)\right]$$

Danskin's theorem helps us compute the gradient of $F_\varepsilon^\odot$:

$$\begin{aligned}
\nabla_\theta F_\varepsilon^\odot(\theta) &= \mathbb{E}\left[e^{-(\varepsilon Z - \varepsilon^2/2)} \odot \nabla_\theta \left(\max_{v \in \mathcal{V}} f_{\varepsilon,Z}(\theta, v)\right)\right] \\
&= \mathbb{E}\left[e^{-(\varepsilon Z - \varepsilon^2/2)} \odot \nabla_1 f_{\varepsilon,Z}\left(\theta, \arg\max_{v \in \mathcal{V}} f_{\varepsilon,Z}(\theta, v)\right)\right] \\
&= \mathbb{E}\left[e^{-(\varepsilon Z - \varepsilon^2/2)} \odot e^{\varepsilon Z - \varepsilon^2/2} \odot \arg\max_{v \in \mathcal{V}} f_{\varepsilon,Z}(\theta, v)\right] \\
&= \mathbb{E}\left[\arg\max_{v \in \mathcal{V}} \left(\theta \odot e^{\varepsilon Z - \varepsilon^2/2}\right)^\top v\right] \\
&= \widehat{f}_\varepsilon^\odot(\theta)
\end{aligned}$$

And we conclude

$$\widehat{f}_\varepsilon^\odot(\theta) = \arg\max_{\mu \in \mathrm{dom}(\Omega_\varepsilon^\odot)} \{\theta^\top \mu - \Omega_\varepsilon^\odot(\mu)\}$$

while keeping in mind that $\mathrm{dom}(\Omega_\varepsilon^\odot) \not\subseteq \mathrm{conv}(\mathcal{V})$.