Analyzing the Variety Loss in the Context of Probabilistic Trajectory Prediction

Luca Anthony Thiede *
Georg-August-Universität Göttingen
luca.thiede@yahoo.com

Pratik Prabhanjan Brahma
Innovation and Engineering Center California
Volkswagen Group of America
pratik.brahma@vw.com

1. Supplementary Material

Proof of Theorem 1

Proof. For the sake of simplicity we only consider the one dimensional case.
First we bin the support of P_T in M equally sized bins $b_1, b_2, ..., b_M$ of width 2ϵ. Then we can write the MoN Loss as

$$L_N(P_T, P) \approx \sum_{i=1}^{M} P_T(b_i) \int_{b_i} \text{EMoN}_{P,b_i}(x^*)\, \mathrm{d}x^* \tag{1}$$

with

$$\text{EMoN}_{P,b_i}(x^*) = \int_{b_i} \min\left(|x^* - x_1|, |x^* - x_2|, ..., |x^* - x_N|\right) P(x_1)P(x_2)\dots P(x_N)\,\mathrm{d}x_1\,\mathrm{d}x_2\dots\mathrm{d}x_N \tag{2}$$

In expectation there are $NP(b_i)$ samples in bin b_i. Assume, that $z_i = NP(b_i)$ is an integer. Denote the sample that fall in b_i as $H_i \coloneqq \{x_1^i, x_2^i, ..., x_n^i\}$. Then, for the calculation of $\text{EMoN}_{P,b_i}(x^*)$ we can ignore all samples that are not in H_i. Then we can write

$$\int_{b_i} \text{EMoN}_{P,b_i}(x^*)\,\mathrm{d}x^* = \tag{3}$$

$$\int_{b_i}\int_{b_i} \min\left(|x^* - x_1^i|, |x^* - x_2^i|, ..., |x^* - x_{z_i}^i|\right) P(x_1^i)P(x_2^i)...P(x_{z_i}^i)\,\mathrm{d}x_1^i\,\mathrm{d}x_2^i, ..., \mathrm{d}x_{z_i}^i\,\mathrm{d}x^* \tag{4}$$

For small ϵ and because we only consider differentiable functions, we can approximate $P(x_1^i)P(x_2^i)...P(x_{z_i}^i)$ withing the bin as uniform. Next we have to calculate the inner integral in (4). For that, center the bins around the origin (which is possible because we have uniform probability distributions) and consider the survival function of individual random samples $|x_i - x^*|$ [1]:

$$S(x) = \Pr\{|x_i - x^*| > x\} \tag{5}$$

$$= \Pr\{x_i > x^* + x\} + \Pr\{x_i < x^* - x\} \tag{6}$$

$$= \begin{cases} 1 & \text{if } x \le 0 \\ 1 - \frac{x}{\epsilon} & \text{if } 0 < x \le \epsilon - |x^*| \\ \frac{\epsilon + |x^*| - x}{2\epsilon} & \text{if } \epsilon - |x^*| < x < \epsilon + |x^*| \\ 0 & \text{if } x \ge \epsilon + |x^*| \end{cases} \tag{7}$$

Then the survival function $S_{MoN,z_i}(x)$ of $\min|x_i - x^*|$ is the probability, that all z_i sample will independently be bigger than x. Thus the survival function is $S_{MoN,z_i}(x) = S(x)^{z_i}$ and therefore

$$E[\min|x_i - x^*|] = \tag{8}$$

$$= \int_0^{+\infty} S(x)^{z_i} dx \tag{9}$$

$$= \int_0^{\epsilon - |x^*|} \left(1 - \frac{x}{\epsilon}\right)^{z_i} dx + \int_{\epsilon - |x^*|}^{\epsilon + |x^*|} \left(\frac{\epsilon + |x^*| - x}{2\epsilon}\right)^{z_i} dx \tag{10}$$

$$= \frac{-\epsilon}{z_i + 1}\left(1 - \frac{x}{\epsilon}\right)^{z_i+1}\Bigg|_0^{\epsilon - |x^*|} + \frac{-2\epsilon}{z_i+1}\left(\frac{\epsilon + |x^*| - x}{2\epsilon}\right)^{z_i+1}\Bigg|_{\epsilon - |x^*|}^{\epsilon + |x^*|} \tag{11}$$

$$= \frac{\epsilon}{z_i + 1}\left(1 - \frac{|x^*|^{z_i+1}}{\epsilon^{z_i+1}} + 2\frac{|x^*|^{z_i+1}}{\epsilon^{z_i+1}}\right) \tag{12}$$

$$= \frac{\epsilon}{z_i + 1}\left(1 + \frac{|x^*|^{z_i+1}}{\epsilon^{z_i+1}}\right) \tag{13}$$

*Work done during internship at Volkswagen Group of America Innovation and Engineering Center located in Belmont, California.

Substituting (13) in (1) therefore yields

$$L_N(P_T, P) \approx \sum_{i=1}^{M} P_T(b_i) \int_{-\epsilon}^{\epsilon} \frac{\epsilon \left(1 + \frac{|x^*|^{z_i+1}}{\epsilon^{z_i+1}}\right)}{z_i + 1} \, \mathrm{d}x^* \tag{14}$$

$$= \epsilon \sum_{i=1}^{M} P_T(b_i) \int_{-\epsilon}^{\epsilon} \frac{\left(1 + \frac{|x^*|^{z_i+1}}{\epsilon^{z_i+1}}\right)}{z_i + 1} \, \mathrm{d}x^* \tag{15}$$

$$= \epsilon \sum_{i=1}^{M} P_T(b_i) \frac{2\epsilon + 2\int_0^{\epsilon} \frac{x^{*\,z_i+1}}{\epsilon^{z_i+1}} \, \mathrm{d}x^*}{z_i + 1} \tag{16}$$

$$= \underbrace{2\epsilon^2}_{:=a} \sum_{i=1}^{M} P_T(b_i) \frac{1 + \frac{1}{NP(b_i)+2}}{NP(b_i) + 1} \tag{17}$$

$$\approx a \sum_{i=1}^{M} P_T(b_i) \frac{1 + \frac{1}{NP(b_i)}}{NP(b_i)} \tag{18}$$

$$= a \sum_{i=1}^{M} P_T(b_i) \left(\frac{1}{NP(b_i)} + \frac{1}{(NP(b_i))^2} \right) \tag{19}$$

$$\approx a \sum_{i=1}^{M} P_T(b_i) \frac{1}{NP(b_i)} \tag{20}$$

where we used in line (18) and (20) that $N \to \infty$. We can find the minimum of $L_N(P_T, P)$ under the constraint that $\sum_{i=1}^{M} P(b_i) = 1$ by using a Lagrange multiplier. The objective is therefore:

$$f(P(b_i), \lambda) := a \sum_{i=1}^{M} \frac{P_T(b_i)}{NP(b_i)} - \lambda \left(1 - \sum_{i=1}^{M} P(b_i)\right) \tag{21}$$

Optimization for $P(b_i)$ yields:

$$\nabla_{P(b_i)} f(P(b_i), \lambda) = \lambda - a \frac{P_T(b_i)}{NP(b_i)^2} \stackrel{!}{=} 0 \tag{22}$$

$$\Leftrightarrow P(b_i) = \sqrt{\frac{aP_T(b_i)}{\lambda N}} \tag{23}$$

$$\nabla_{\lambda} f(P(b_i), \lambda) = 1 - \sum_{i=1}^{M} P(b_i) \stackrel{!}{=} 0 \tag{24}$$

$$\Leftrightarrow \sum_{i=1}^{M} P(b_i) = 1 \tag{25}$$

where we can omit the $\pm$ in (23), because probabilites can not be negative. Putting (23) and (25) together results in:

$$\Rightarrow \sum_{i=1}^{M} P_T(b_i) \left(\sqrt{\frac{aP_T(b_i)}{\lambda N}} \right) = 1 \tag{26}$$

$$\Rightarrow \lambda = \left(\frac{\sum_{i=1}^{M} \sqrt{aP_T(b_i)}}{\sqrt{N}} \right)^2 \tag{27}$$

$$\Rightarrow P(b_i) = \frac{\sqrt{P_T(b_i)}}{\sum_{i=1}^{M} \sqrt{P_T(b_i)}} \tag{28}$$

Note, that this result is true for any distance metric in any dimensions, as long as the expected minimum distance of N sample to a target sample goes locally with $\mathcal{O}(\frac{1}{N})$. $\square$

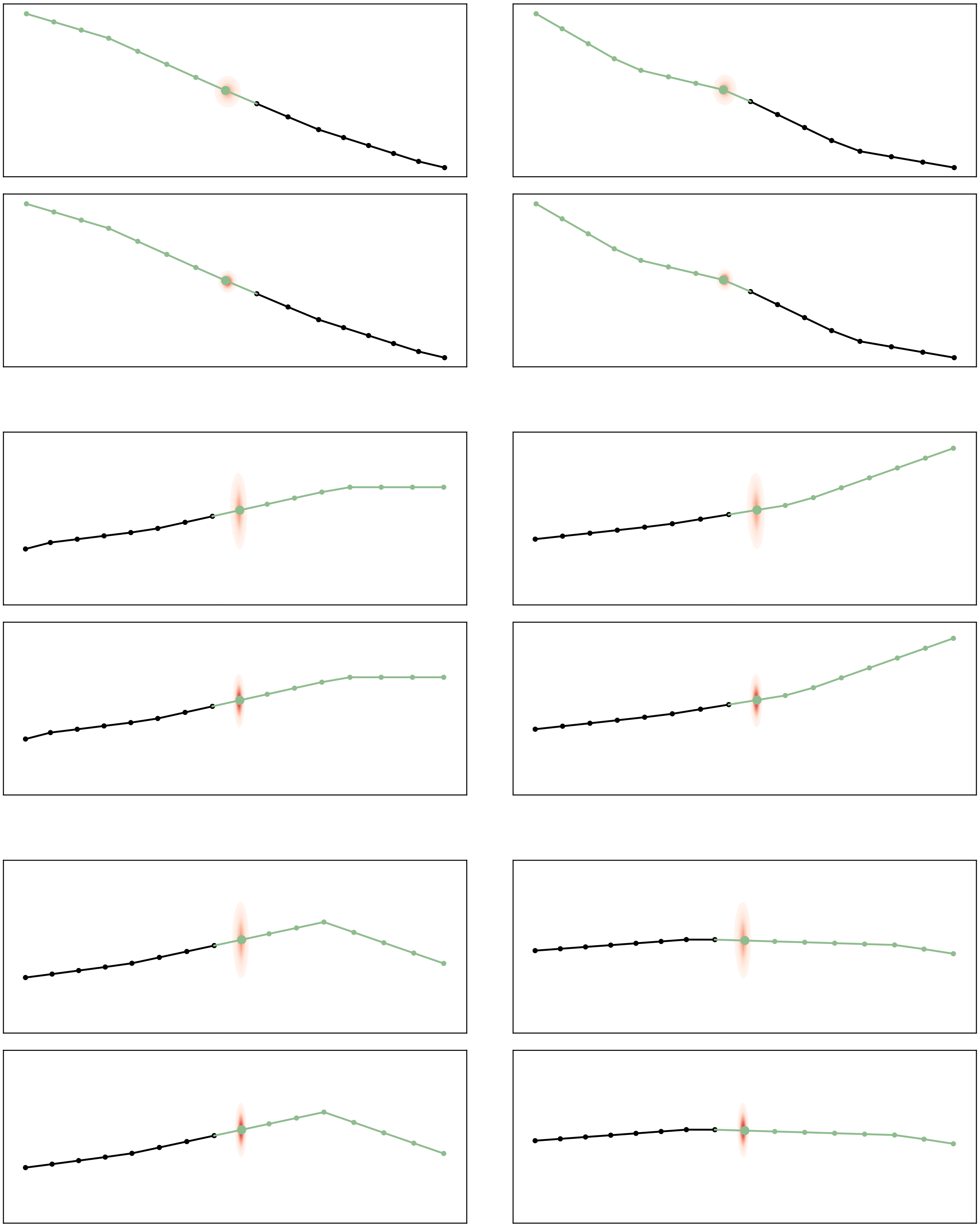

Figure 1: Marginalized probabilities of Social–GAN on the Zara dataset for $t = 1$. Upper panel is the uncompensated PDF and the lower panel is the compensated one. Black is the observed trajectory and green the ground truth future.

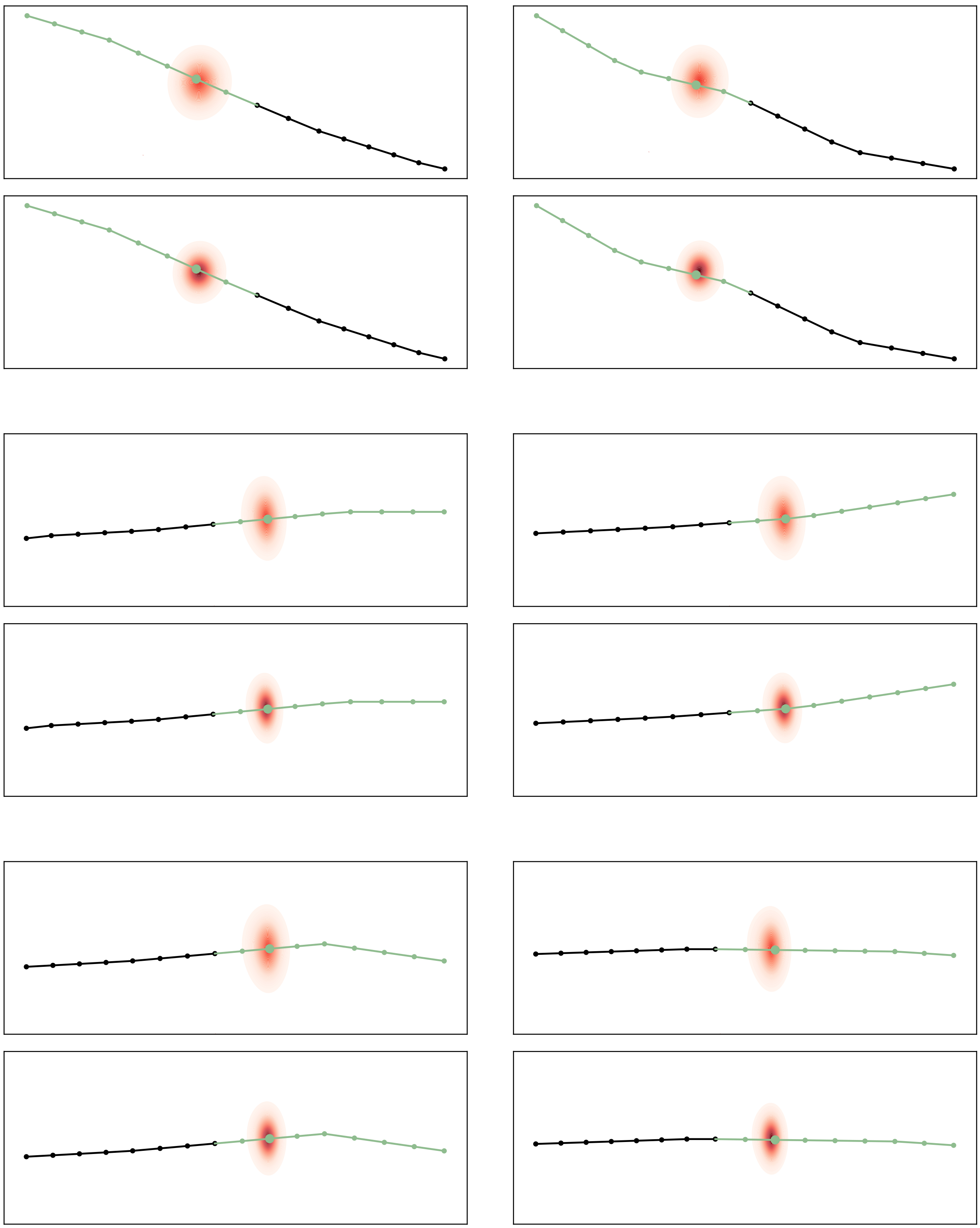

Figure 2: Marginalized probabilities of Social–GAN on the Zara dataset for $t = 2$. Upper panel is the uncompensated PDF and the lower panel is the compensated one. Black is the observed trajectory and green the ground truth future.

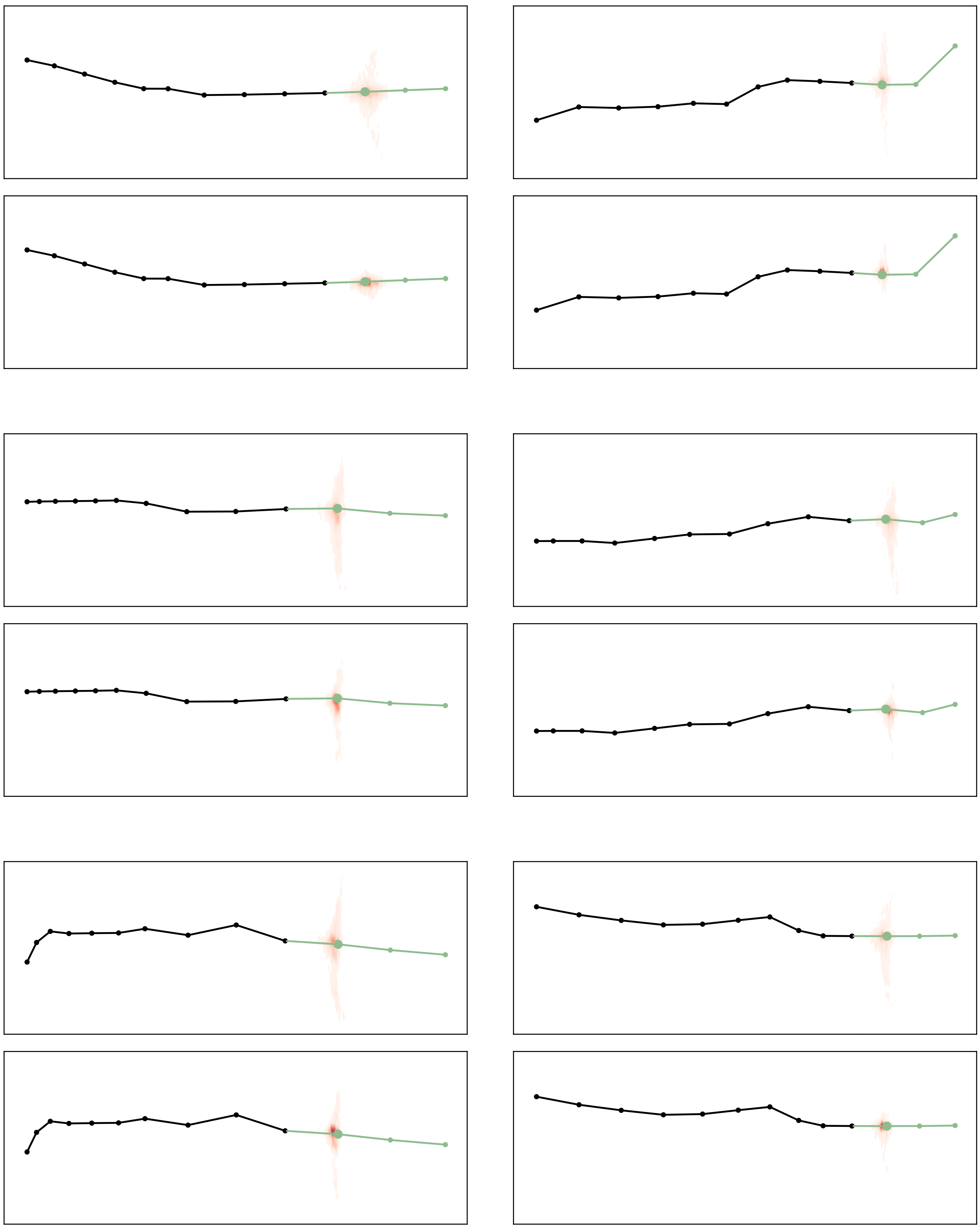

Figure 3: Marginalized probabilities of our own model on the NGSIM dataset for $t = 1$. Upper panel is the uncompensated PDF and the lower panel is the compensated one. Black is the observed trajectory and green the ground truth future.

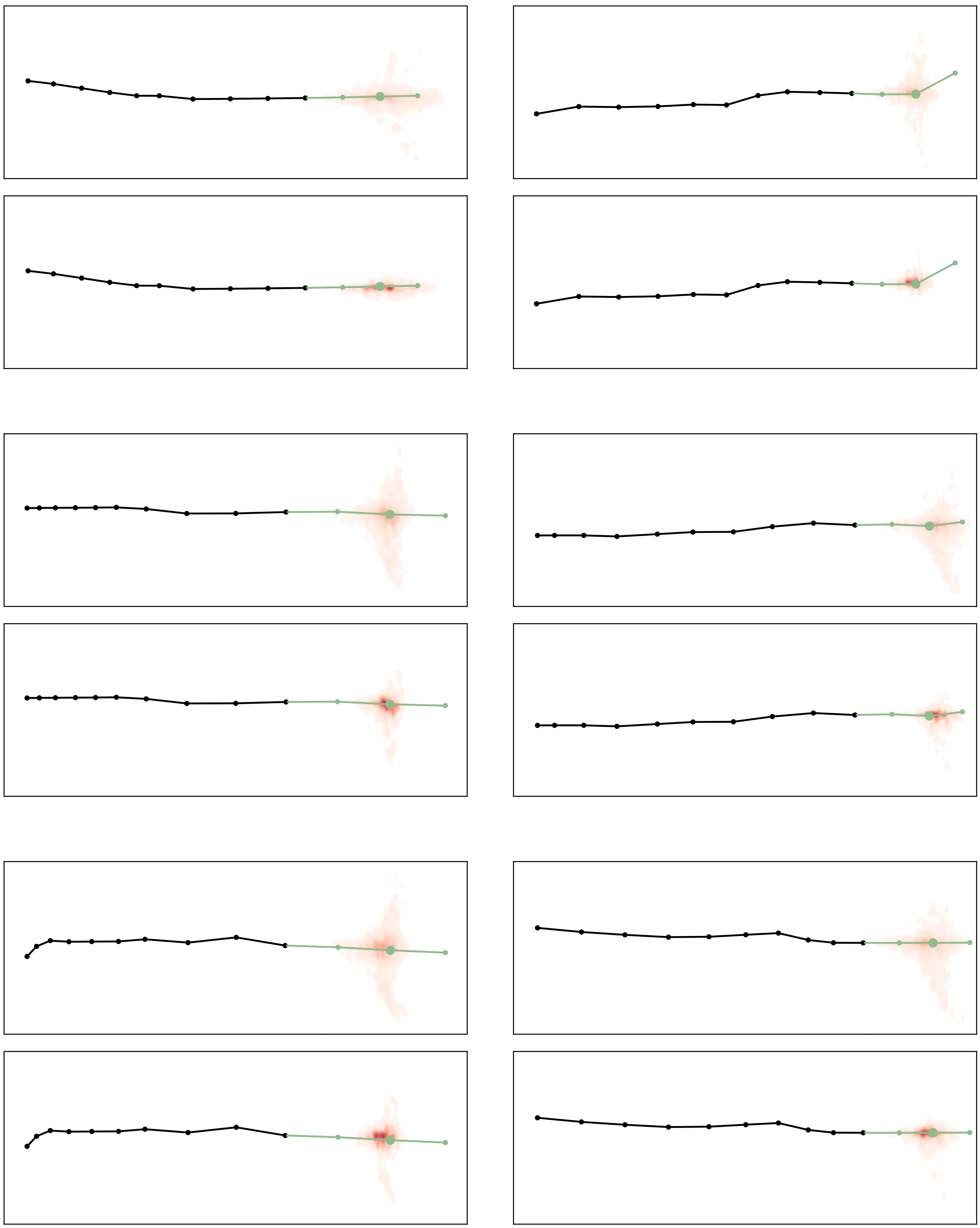

Figure 4: Marginalized probabilities of our own model on the NGSIM dataset for $t = 2$. Upper panel is the uncompensated PDF and the lower panel is the compensated one. Black is the observed trajectory and green the ground truth future.

References

[1] BGM (https://math.stackexchange.com/users/297308/bgm). Expected minimum absolute difference to a given point correctly computed? Mathematics Stack Exchange. URL:https://math.stackexchange.com/q/3000933 (version: 2018-11-16).

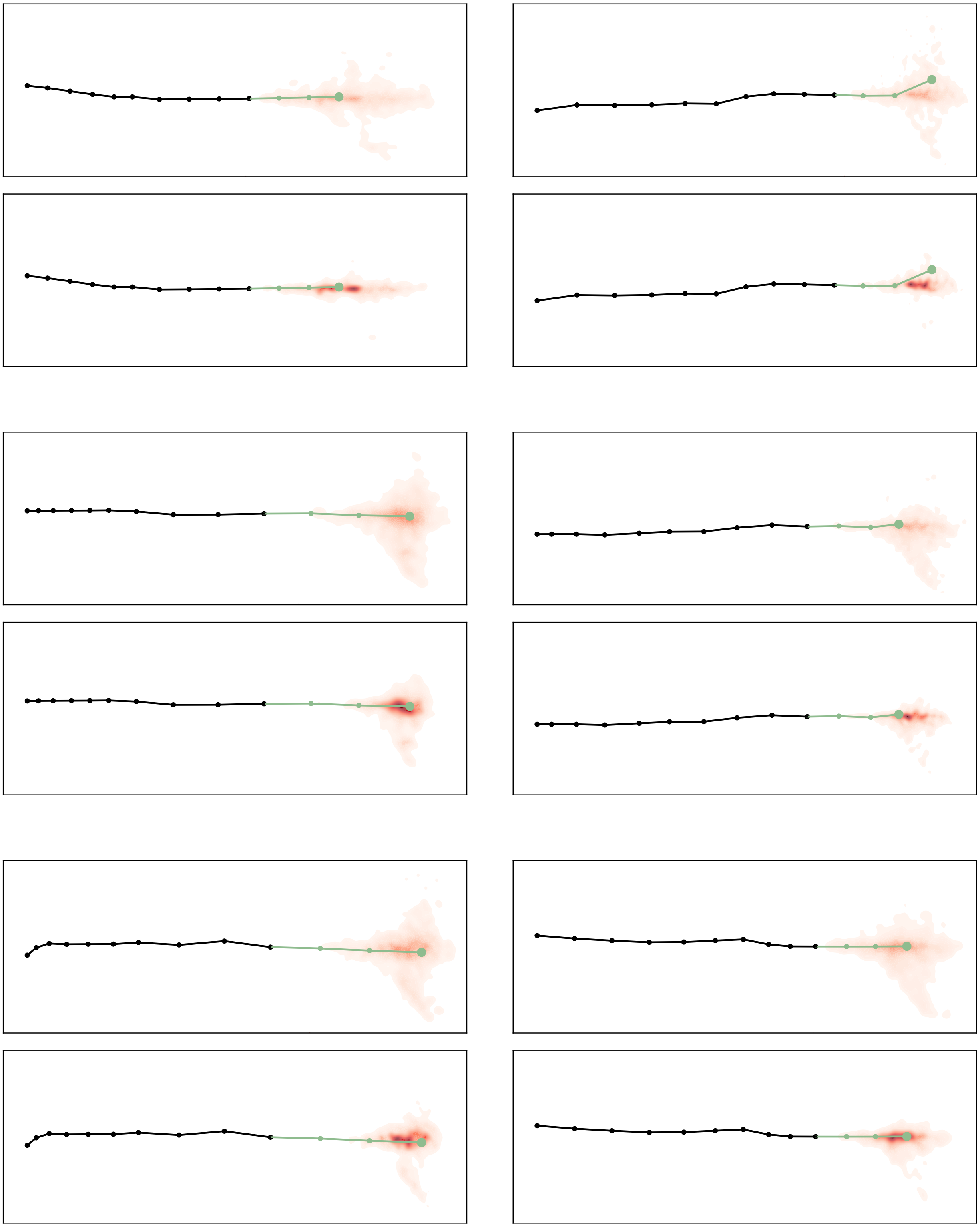

Figure 5: Marginalized probabilities of our own model on the NGSIM dataset for $t = 3$. Upper panel is the uncompensated PDF and the lower panel is the compensated one. Black is the observed trajectory and green the ground truth future.