Complement of proof for Proposition 6 from my article "Coarse Ricci curvature of Markov chains on metric spaces".

In the notation of Section 8, we have to prove that

$$W_1(\mu_0, \mu_1) = d(x,y)\left(1 - \frac{\varepsilon^2}{2(N+2)}\operatorname{Ric}(v,v) + O(\varepsilon^3 + \delta\varepsilon^2)\right) \tag{1}$$

where μ_0 and μ_1 are the uniform measures on the ε-balls around x and y respectively.

Let $\mathrm{d}w$ be the uniform probability measure over w in the unit ball of the tangent space at x. Let μ'_0 be, as in the article, the image of this measure under $w \mapsto \exp_x(\varepsilon w)$, and μ'_1 the image of this measure under $w \mapsto \exp_y(\varepsilon w)$ (where we identify the tangent spaces at x and y via parallel transport along the geodesic $\exp_x(\delta v)$; this preserves the uniform measure on the unit ball because parallel transport is isometric).

The density of μ'_0 and μ'_1 wrt μ_0 and μ_1 is $1 + O(\varepsilon^2)$ by properties of the exponential map. More precisely, defining likewise μ_z and μ'_z for any point z in the manifold (in a neighborhood of x), expand $\frac{\mathrm{d}\mu_z}{\mathrm{d}\mu'_z}(w) = 1 + \varepsilon^2 f_z(\varepsilon, w)$ (where we identify w in tangent spaces at z to the tangent space at x via parallel transport for z in some neighborhood of x). Since $y = \exp_x(\delta v)$, we can expand again $f_y(\varepsilon, w) = f_x(\varepsilon, w) + \delta g(\delta, \varepsilon, w)$. Both f and g are given by various derivatives of the exponential map. Note that $\int_w f_x(\varepsilon, w)\,\mathrm{d}w = 0$ since both μ_x and μ'_x are probability measures.

We have to estimate the W_1 distance between the images by $\exp(\varepsilon w)$ of the measures $(1 + \varepsilon^2 f_x(\varepsilon, w))\,\mathrm{d}w$ and $(1 + \varepsilon^2 f_x(\varepsilon, w) + \delta\varepsilon^2 g(\delta, \varepsilon, w))\,\mathrm{d}w$, where the exponential is based at x and y respectively, namely, letting $\varphi_z : w \mapsto \exp_z(\varepsilon w)$, we have to compute

$$W_1\left(\varphi_x^*\left((1 + \varepsilon^2 f_x(\varepsilon, w))\,\mathrm{d}w\right), \varphi_y^*\left((1 + \varepsilon^2 f_x(\varepsilon, w) + \delta\varepsilon^2 g(\delta, \varepsilon, w))\,\mathrm{d}w\right)\right) \tag{2}$$

We apply the triangle inequality to estimate separately

$$W_1\left(\varphi_x^*\left((1 + \varepsilon^2 f_x(\varepsilon, w))\,\mathrm{d}w\right), \varphi_y^*\left((1 + \varepsilon^2 f_x(\varepsilon, w))\,\mathrm{d}w\right)\right) \tag{3}$$

and

$$W_1\left(\varphi_y^*\left((1 + \varepsilon^2 f_x(\varepsilon, w))\,\mathrm{d}w\right), \varphi_y^*\left((1 + \varepsilon^2 f_x(\varepsilon, w) + \delta\varepsilon^2 g(\delta, \varepsilon, w))\,\mathrm{d}w\right)\right) \tag{4}$$

By Kantorovich duality, the second one is equal to $\sup_h \int_w h(\varphi_y(w))\,\delta\varepsilon^2 g(\delta, \varepsilon, w)\,\mathrm{d}w$ where the supremum is on all 1-Lipshitz functions h. By shifting h we can assume that $h(y) = 0$ (because g integrates to 0 because all measures are probability measures). Then $|h(\varphi_y(w))| \leqslant \varepsilon$ since $\varphi_y(w)$ is in the ε-ball around y. So the corresponding term is $O(\delta\varepsilon^3)$.

We can bound the first term by choosing a specific coupling, for instance, the trivial coupling, for which $d(\varphi_x(w), \varphi_y(w)) = \delta(1 - \varepsilon^2 K(v,w)/2 + O(\varepsilon^3 +$

$\delta\varepsilon^2$)) by Proposition 6. We have to integrate this against $(1+\varepsilon^2 f_x(\varepsilon,w))\,\mathrm{d}w$, this yields

$$\int_w \delta(1-\varepsilon^2 K(v,w)/2 + O(\varepsilon^3+\delta\varepsilon^2))(1+\varepsilon^2 f_x(\varepsilon,w))\,\mathrm{d}w \tag{5}$$

$$= \delta - \frac{\delta\varepsilon^2}{2}\int_w K(v,w)\,\mathrm{d}w + \delta\varepsilon^2\int_w f_x(\varepsilon,w)\,\mathrm{d}w + \delta\, O(\varepsilon^3+\delta\varepsilon^2) \tag{6}$$

and we conclude using $\int_w f_x(\varepsilon,w)\,\mathrm{d}w = 0$ and $\int_w K(v,w)\,\mathrm{d}w = \frac{1}{N+2}\mathrm{Ric}(v,v)$.