Bag of Results

Lemma .1.

$$q_t = \mathcal{Q}\left(w_t + \frac{X_t^T u_{t-1}}{\|X_t\|_2^2}\right).$$

Proof. This follows simply by completing a square. Provided $X_t \neq 0$, we have by the definition of q_t

$$\begin{aligned} q_t &= \operatorname*{argmin}_{p\in\{-1,0,1\}} \|u_{t-1} + (w_t - p)X_t\|_2^2 = \operatorname*{argmin}_{p\in\{-1,0,1\}} (w_t - p)^2 + 2(w_t - p)\frac{X_t^T u_{t-1}}{\|X_{t-1}\|_2^2} \\ &= \operatorname*{argmin}_{p\in\{-1,0,1\}} \left((w_t - p) + \frac{X_t^T u_{t-1}}{\|X_{t-1}\|_2^2}\right)^2 - \left(\frac{X_t^T u_{t-1}}{\|X_{t-1}\|_2^2}\right)^2. \end{aligned}$$

Because the former term is always non-negative, it must be the case that the minimizer is $\mathcal{Q}\left(w_t + \frac{X_t^T u_{t-1}}{\|X_t\|_2^2}\right)$. □

Lemma .2. *Suppose* $0 < w_t < 1$*, and* X_t *a random vector in* R^d*. Then*

$$\mathbb{P}_{X_t}\left((w_t - q_t)^2 + 2(w_t - q_t)X_t^T u_{t-1} > \alpha \middle| u_{t-1}\right) = \begin{cases} \mu_y\left(\frac{\alpha-(w_t+1)^2}{2(w_t+1)}, \frac{\alpha-(w_t-1)^2}{2(w_t-1)}\right) & \alpha < -w_t - w_t^2 \\ \mu_y\left(\frac{\alpha-w_t^2}{2w_t}, \frac{\alpha-(w_t-1)^2}{2(w_t-1)}\right) & -w_t - w_t^2 \leq \alpha \leq w_t - w_t^2 \\ 0 & \alpha > w_t - w_t^2 \end{cases}$$

where μ_y *is the probability measure over* $\mathbb{R}$ *induced by the random variable* $y = X_t^T u_{t-1}$*.*

Proof. Let A_b denote the event that $q_t = b$ for $b \in \{-1, 0, 1\}$. Then by the law of total probability

$$\mathbb{P}\left((w_t - q_t)^2 + 2(w_t - q_t)y > \alpha \middle| u_{t-1}\right) = \sum_{b\in\{-1,0,1\}} \mathbb{P}\left((w_t - b)^2 + 2(w_t - b)y > \alpha \middle| u_{t-1}, A_b\right).$$

- **b = 0** $q_t = 0$ precisely when $-1/2 - w_t \leq y \leq 1/2 - w_t$. So we have

$$\begin{aligned} &\mathbb{P}\left(w_t^2 + 2w_t y > \alpha \middle| u_{t-1}, A_b\right) = \mathbb{P}\left(y > \frac{\alpha - w_t^2}{2w_t} \middle| u_{t-1}, -1/2 - w_t \leq y \leq 1/2 - w_t\right) \\ &= \begin{cases} \mu_y(-1/2 - w_t, 1/2 - w_t) & \alpha < -w_t - w_t^2 \\ \mu_y\left(\frac{\alpha-w_t^2}{2w_t}, 1/2 - w_t\right) & -w_t - w_t^2 \leq \alpha \leq w_t - w_t^2 \\ 0 & \alpha > w_t - w_t^2 \end{cases}. \end{aligned}$$

- **b = 1** $q_t = 1$ precisely when $y > 1/2 - w_t$. Noting that $w_t - 1 < 0$, we have

$$\begin{aligned} &\mathbb{P}\left((w_t - 1)^2 + 2(w_t - 1)y > \alpha \middle| u_{t-1}, A_1\right) = \mathbb{P}\left(y < \frac{\alpha - (w_t - 1)^2}{2(w_t - 1)} \middle| u_{t-1}, y > 1/2 - w_t\right) \\ &= \begin{cases} \mu_y\left(1/2 - w_t, \frac{\alpha-(w_t-1)^2}{2(w_t-1)}\right) & \alpha \leq w_t - w_t^2 \\ 0 & \alpha > w_t - w_t^2 \end{cases}. \end{aligned}$$

- **b = -1** $q_t = -1$ precisely when $y < -1/2 - w_t$. So we have

$$\begin{aligned} &\mathbb{P}\left((w_t + 1)^2 + 2(w_t - 1)y > \alpha \middle| u_{t-1}, A_1\right) = \mathbb{P}\left(y > \frac{\alpha - (w_t + 1)^2}{2(w_t + 1)} \middle| u_{t-1}, y < -1/2 - w_t\right) \\ &= \begin{cases} \mu_y\left(\frac{\alpha-(w_t+1)^2}{2(w_t+1)}, -1/2 - w_t\right) & \alpha \leq -w_t - w_t^2 \\ 0 & \alpha > -w_t - w_t^2 \end{cases}. \end{aligned}$$

Summing these three functions yields the result. □

Corollary .3. *When $-1 < w_t < 0$, we have*

$$\mathbb{P}_{X_t}\left((w_t-q_t)^2+2(w_t-q_t)X_t^T u_{t-1} > \alpha \middle| u_{t-1}\right) = \begin{cases} \mu_y\left(\frac{\alpha-(w_t+1)^2}{2(w_t+1)}, \frac{\alpha-(w_t-1)^2}{2(w_t-1)}\right) & \alpha < w_t - w_t^2 \\ \mu_y\left(\frac{\alpha-(w_t+1)^2}{2(w_t+1)}, \frac{\alpha-w_t^2}{2w_t}\right) & w_t - w_t^2 \le \alpha \le -w_t - w_t^2 \\ 0 & \alpha > w_t - w_t^2 \end{cases}.$$

Lemma .4. *Let $X_t \sim \mathrm{Unif}(S^{d-1})$. Then*

$$\mathbb{E}[q_t X_t | u_{t-1}] = \frac{\mathrm{Area}(S^{d-2})}{(d-1)\cdot \mathrm{Area}(S^{d-1})}\left(\left(1-\min\left\{\left(\frac{1/2-w_t}{\|u_{t-1}\|}\right)^2, 1\right\}\right)^{\frac{d-1}{2}} + \left(1-\min\left\{\left(\frac{1/2+w_t}{\|u_{t-1}\|}\right)^2, 1\right\}\right)^{\frac{d-1}{2}}\right)\frac{u_{t-1}}{\|u_{t-1}\|_2},$$

where μ_y is the probability measure over $\mathbb{R}$ induced by the random variable $y = X_t^T u_{t-1}$.

Proof. By unitary invariance of the X_t, we may assume that $u_{t-1} = \|u_{t-1}\|_2 e_1$. Under this assumption, we have

$$\mathbb{E}[q_t X_t | u_{t-1}] = \mathbb{E}[\mathcal{Q}\left(w_t + \|u_{t-1}\|_2 X_{t,1}\right) X_t | u_{t-1}].$$

Breaking this apart into the events $A_{-1} := \{X_{t,1} \le \frac{-1/2-w_t}{\|u_{t-1}\|}\}$ and $A_1 := \{X_{t,1} \ge \frac{1/2-w_t}{\|u_{t-1}\|}\}$, we note that these regions on S^{d-1} are symmetric about the $u_{t-1} = e_1$ axis. So,

$$\mathbb{E}[q_t X_t | u_{t-1}] = \mathbb{E}[X_t | u_{t-1}, A_1] - \mathbb{E}[X_t | u_{t-1}, A_{-1}].$$

Using the hyperspherical coordinates

$$f\left(\begin{bmatrix}\varphi_1\\ \varphi_2\\ \vdots\\ \varphi_{d-1}\end{bmatrix}\right) = \begin{bmatrix}\cos(\varphi_1)\\ \sin(\varphi_1)\cos(\varphi_2)\\ \sin(\varphi_1)\sin(\varphi_2)\cos(\varphi_3)\\ \vdots\\ \sin(\varphi_1)\dots\sin(\varphi_{d-2})\cos(\varphi_{d-1})\\ \sin(\varphi_1)\dots\sin(\varphi_{d-2})\sin(\varphi_{d-1})\end{bmatrix}, \tag{1}$$

we have $\mathbb{E}[X_{t,1}|u_{t-1}, A_1] = 0$ if $\frac{1/2-w_t}{\|u_{t-1}\|_2} > 1$. Otherwise,

$$\begin{aligned}\mathbb{E}[X_{t,1}|u_{t-1}, A_1] &= \frac{1}{\mathrm{Area}(S^{d-1})}\int_0^{\arccos(\frac{1/2-w_t}{\|u_{t-1}\|_2})} \cos(\varphi_1)\, dA \\ &= \frac{\mathrm{Area}(S^{d-2})}{\mathrm{Area}(S^{d-1})}\int_0^{\arccos(\frac{1/2-w_t}{\|u_{t-1}\|_2})} \cos(\varphi_1)\sin^{d-2}(\varphi_1)\, d\varphi_1 = \frac{\mathrm{Area}(S^{d-2})}{(d-1)\mathrm{Area}(S^{d-1})}\left(1-\left(\frac{1/2-w_t}{\|u_{t-1}\|_2}\right)^2\right)^{\frac{d-1}{2}}.\end{aligned}$$

Similarly, $\mathbb{E}[X_{t,1}|u_{t-1}, A_1] = 0$ if $\frac{1/2-w_t}{\|u_{t-1}\|_2} \le -1$. Otherwise,

$$\mathbb{E}[X_{t,1}|u_{t-1}, A_1] = \frac{\mathrm{Area}(S^{d-2})}{(d-1)\mathrm{Area}(S^{d-1})}\left(1-\left(\frac{1/2+w_t}{\|u_{t-1}\|_2}\right)^2\right)^{\frac{d-1}{2}}.$$

□

Lemma .5. *Let $d \ge 2$. Then*

$$\sqrt{\frac{d}{4e^3}}. \le \frac{\mathrm{Area}(S^{d-2})}{\mathrm{Area}(S^{d-1})} \le \frac{e}{\pi}\sqrt{d}$$

Proof. Recall that the Stirling bounds on the Gamma function give

$$\sqrt{2\pi}\, n^{n+1/2} e^{-n} \le \Gamma(n) \le e n^{n+1/2} e^{-n}$$

for all $n \in \mathbb{N}$. Using these bounds along with the fact that $\text{Area}(S^{d-1}) = \frac{d\pi^{d/2}}{\Gamma\left(\frac{d}{2}+1\right)}$, we have for the lower bound [**EL Note:** *Of course, it can never be the case that both $\frac{d}{2}$ and $\frac{d-1}{2}$ are both integers. This seems like a small, annoying detail though.*]

$$\frac{\text{Area}(S^{d-2})}{\text{Area}(S^{d-1})} = \frac{d-1}{d}\frac{1}{\sqrt{\pi}}\frac{\Gamma\left(\frac{d}{2}+1\right)}{\Gamma\left(\frac{d-1}{2}+1\right)} \geq \frac{1}{2}\frac{\sqrt{2\pi}}{e\sqrt{\pi}}\frac{(\frac{d+2}{2})^{\frac{d+3}{2}}e^{-\frac{d+2}{2}}}{(\frac{d+1}{2})^{\frac{d+2}{2}}e^{\frac{-d+1}{2}}} = \frac{1}{\sqrt{2e^3}}\frac{(\frac{d+2}{2})^{\frac{d+3}{2}}}{(\frac{d+1}{2})^{\frac{d+2}{2}}}$$
$$= \sqrt{\frac{d+2}{4e^3}}\left(\frac{d+2}{d+1}\right)^{\frac{d+2}{2}} \geq \sqrt{\frac{d}{4e^3}}.$$

Now, for the upper bound,

$$\frac{\text{Area}(S^{d-2})}{\text{Area}(S^{d-1})} = \frac{d-1}{d}\frac{1}{\sqrt{\pi}}\frac{\Gamma\left(\frac{d}{2}+1\right)}{\Gamma\left(\frac{d-1}{2}+1\right)} \leq \frac{1}{\sqrt{\pi}}\frac{e\left(\frac{d+2}{2}\right)^{\frac{d+3}{2}}e^{-\frac{d+2}{2}}}{\sqrt{2\pi}\left(\frac{d+1}{2}\right)^{\frac{d+2}{2}}e^{-\frac{d+1}{2}}} = \frac{\sqrt{e}}{\pi\sqrt{2}}\left(\frac{d+2}{2}\right)^{\frac{1}{2}}\left(\frac{d+2}{d+1}\right)^{\frac{d+2}{2}}$$
$$\leq \frac{\sqrt{e}}{\pi\sqrt{2}}\sqrt{d}\left(1+\frac{1}{d+1}\right)^{\frac{d+1}{2}}\cdot\left(1+\frac{1}{d+1}\right)^{\frac{1}{2}} \leq \frac{e}{\pi}\sqrt{d}.$$

□

Lemma .6. *For any $\varepsilon > 0$ and $C_\varepsilon = \frac{(4+\varepsilon)e^{21/8}}{2}$, we have*

$$\mathbb{E}\left[\|u_t\|_2^2 - \|u_{t-1}\|_2^2 \middle| \|u_{t-1}\|_2^2 > C^2(d-1)\right] \leq -\varepsilon.$$

Proof. Let $\varepsilon > 0$. For any fixed u_{t-1} with $\|u_{t-1}\|_2 > C_\varepsilon\sqrt{d-1}$, we have by Lemma .4 and Lemma .5

$$\mathbb{E}\left[\|u_t\|_2^2 - \|u_{t-1}\|_2^2 \middle| u_{t-1}\right] = \mathbb{E}\left[(w_t - q_t)^2 + 2(w_t - q_t)X_t^T u_{t-1} \middle| u_{t-1}\right] \leq 4 - 2\mathbb{E}\left[q_t X_t^T u_{t-1}\middle| u_{t-1}\right]$$
$$= 4 - 2\|u_{t-1}\|_2\frac{\text{Area}(S^{d-2})}{(d-1)\cdot\text{Area}(S^{d-1})}\left(\left(1-\left(\frac{1/2-w_t}{\|u_{t-1}\|}\right)^2\right)^{(d-1)/2} + \left(1-\left(\frac{1/2+w_t}{\|u_{t-1}\|}\right)^2\right)^{(d-1)/2}\right)$$
$$\leq 4 - 2C_\varepsilon\frac{\text{Area}(S^{d-2})}{\sqrt{d}\text{Area}(S^{d-1})}\left(\left(1-\frac{(1/2-w_t)^2}{d-1}\right)^{(d-1)/2} + \left(1-\frac{(1/2+w_t)^2}{d}\right)^{(d-1)/2}\right)$$
$$\leq 4 - 2C_\varepsilon\frac{\text{Area}(S^{d-2})}{\sqrt{d}\text{Area}(S^{d-1})}\left(\exp\left(-\frac{(1/2-w_t)^2}{2}\right) + \exp\left(-\frac{(1/2+w_t)^2}{2}\right)\right)$$
$$\leq 4 - 2C_\varepsilon\sqrt{\frac{1}{4e^3}}\left(\exp\left(-\frac{(1/2-w_t)^2}{2}\right) + \exp\left(-\frac{(1/2+w_t)^2}{2}\right)\right)$$
$$\leq 4 - 4C_\varepsilon e^{-\frac{9}{8}}\sqrt{\frac{1}{4e^3}} = 4 - \frac{4C_\varepsilon}{2e^{21/8}} = -\varepsilon < 0.$$

□

Lemma .7. *Let w_t be fixed. If $X_t \sim \text{Unif}(S^{d-1})$, then for any $\lambda \in (0,1)$ and $\beta > 2$*

$$\mathbb{E}\left[e^{\lambda\Delta\|u_t\|_2^2}\middle| \|u_{t-1}\|_2 \geq \beta\right] \leq \frac{e\sqrt{d}}{\beta\pi}\cdot\frac{\max\left\{e^{w_t^2}, e^{(1-w_t)^2}, e^{(1+w_t)^2}\right\}^\lambda}{2\lambda\min\{w_t, 1-w_t, 1+w_t\}}\left(e^{-2\lambda(1-w_t)(1/2-w_t)} + e^{-2\lambda(w_t+1)(1/2+w_t)} + e^{2\lambda w_t(1/2-w_t)}\right).$$

Proof. As we usually have done, we'll split the conditional expectation up into the three events which determine the values of q. Note that whenever $\|u_{t-1}\| \geq 2$, each event has non-zero probability of occurring regardless of w_t (this is why we assume $\beta > 2$). By rotational invariance, we may assume that $u_{t-1} = \|u_{t-1}\|_2 e_1$. Using the hyperspherical coordinates as in (1), we have

$$\begin{aligned}\mathbb{E}\left[e^{\lambda \Delta \|u_t\|_2^2} \middle| \|u_{t-1}\|_2 \geq \beta\right] = \frac{\text{Area}(S^{d-2})}{\text{Area}(S^{d-1})}\Big(& e^{\lambda(w_t-1)^2} \int_{\arccos(1)}^{\arccos\left(\frac{1/2-w_t}{\|u_{t-1}\|}\right)} e^{2\lambda(w_t-1)\|u_{t-1}\|_2 \cos(\varphi)} \sin^{d-2}(\varphi)\, d\varphi \\ &+ e^{\lambda w_t^2} \int_{\arccos\left(\frac{1/2-w_t}{\|u_{t-1}\|_2}\right)}^{\arccos\left(\frac{-1/2-w_t}{\|u_{t-1}\|}\right)} e^{2\lambda w_t \|u_{t-1}\| \cos(\varphi)} \sin^{d-2}(\varphi)\, d\varphi \\ &+ e^{\lambda(w_t+1)^2} \int_{\arccos\left(\frac{-1/2-w_t}{\|u_{t-1}\|_2}\right)}^{\arccos(-1)} e^{2\lambda(w_t+1)\|u_{t-1}\|_2 \cos(\varphi)} \sin^{d-2}(\varphi)\, d\varphi\Big).\end{aligned}$$

We'll control each integral individually. Note for the first in the sum of three, we may decompose it as

$$\begin{aligned}&\int_{\arccos(1)}^{\arccos\left(\frac{1/2-w_t}{\|u_{t-1}\|}\right)} e^{2\lambda(w_t-1)\|u_{t-1}\|_2 \cos(\varphi)} \sin^{d-2}(\varphi)\, d\varphi \\ &= \int_0^{\frac{\pi}{2}} e^{2\lambda(w_t-1)\|u_{t-1}\|_2 \cos(\varphi)} \sin^{d-2}(\varphi)\, d\varphi + \int_{\frac{\pi}{2}}^{\arccos\left(\frac{1/2-w_t}{\|u_{t-1}\|}\right)} e^{2\lambda(w_t-1)\|u_{t-1}\|_2 \cos(\varphi)} \sin^{d-2}(\varphi)\, d\varphi \\ &\leq \int_0^{\frac{\pi}{2}} \sin^{d-2}(\varphi)\, d\varphi + \left|\arccos\left(\frac{1/2-w_t}{\|u_{t-1}\|}\right) - \arccos(0)\right| \max_{\varphi \in \left[\pi/2, \arccos\left(\frac{1/2-w_t}{\|u_{t-1}\|_2}\right)\right]} e^{2\lambda(w_t-1)\|u_{t-1}\|_2 \cos(\varphi)} \\ &\leq \int_0^{\frac{\pi}{2}} \sin^{d-2}(\varphi)\, d\varphi + \max_{|x| \leq \frac{1/2-w_t}{\|u_{t-1}\|_2}} \left|\frac{1}{\sqrt{1-x^2}}\right| \frac{1}{\|u_{t-1}\|_2} \max\left\{e^{2\lambda(w_t-1)(1/2-w_t)}, 1\right\} \\ &\leq \frac{\text{Area}(S^{d-1})}{2\text{Area}(S^{d-2})} + \frac{16}{7\|u_{t-1}\|_2} \max\left\{e^{2\lambda(w_t-1)(1/2-w_t)}, 1\right\} \leq \frac{\text{Area}(S^{d-1})}{2\text{Area}(S^{d-2})} + \frac{16}{7\beta} \max\left\{e^{2\lambda(w_t-1)(1/2-w_t)}, 1\right\}.\end{aligned}$$

The third integral follows a similar argument, namely

$$\begin{aligned}&\int_{\arccos\left(\frac{-1/2-w_t}{\|u_{t-1}\|_2}\right)}^{\arccos(-1)} e^{2\lambda(w_t+1)\|u_{t-1}\|_2 \cos(\varphi)} \sin^{d-2}(\varphi)\, d\varphi \\ &\leq \int_{\arccos\left(\frac{-1/2-w_t}{\|u_{t-1}\|_2}\right)}^{\frac{\pi}{2}} e^{2\lambda(w_t+1)\|u_{t-1}\|_2 \cos(\varphi)} \sin^{d-2}(\varphi)\, d\varphi + \int_{\frac{\pi}{2}}^{\pi} \sin^{d-2}(\varphi)\, d\varphi \\ &\leq \frac{16}{7\beta} \max\left\{e^{2\lambda(w_t-1)(-1/2-w_t)}, 1\right\} + \frac{\text{Area}(S^{d-1})}{2\text{Area}(S^{d-2})}.\end{aligned}$$

Finally, for the second integral we use the same upper bound + mean value theorem trick as we've used twice above

$$\begin{aligned}&\int_{\arccos\left(\frac{1/2-w_t}{\|u_{t-1}\|_2}\right)}^{\arccos\left(\frac{-1/2-w_t}{\|u_{t-1}\|}\right)} e^{2\lambda w_t \|u_{t-1}\| \cos(\varphi)} \sin^{d-2}(\varphi)\, d\varphi \\ &= \int_{\arccos\left(\frac{1/2-w_t}{\|u_{t-1}\|_2}\right)}^{\frac{\pi}{2}} e^{2\lambda w_t \|u_{t-1}\| \cos(\varphi)} \sin^{d-2}(\varphi)\, d\varphi + \int_{\frac{\pi}{2}}^{\arccos\left(\frac{-1/2-w_t}{\|u_{t-1}\|}\right)} e^{2\lambda w_t \|u_{t-1}\| \cos(\varphi)} \sin^{d-2}(\varphi)\, d\varphi \\ &\leq \frac{32}{7\beta} \max\left\{e^{2\lambda w_t(1/2-w_t)}, e^{2\lambda w_t(-1/2-w_t)}, 1\right\}.\end{aligned}$$

Combining these three upper bounds give

$$\mathbb{E}\left[e^{\lambda \Delta \|u_t\|_2^2} \middle| \|u_{t-1}\|_2 \geq \beta\right] \leq \max\left\{e^{w_t^2}, e^{(1+w_t)^2}\right\} \frac{\text{Area}(S^{d-2})}{\text{Area}(S^{d-1})} \left(\frac{\text{Area}(S^{d-1})}{\text{Area}(S^{d-2})}\right.$$

[**EL Note:** *This new proof also seems to be very sensitive to how close w_t is to ± 1. If you can't modify this version to get it to work, uncomment out the block below and use it instead. On second thought, this is entirely unsurprising. $\Delta\|u_{t-1}\|_2^2 = 0$ if $w_t = 0, \pm 1$, so the m.g.f can't be bounded away from the origin in these cases.*]

∎

Theorem .8. *For any $t \in \mathbb{N}$, we have*

$$\mathbb{P}\left(\|u_t\|_2^2 > \alpha\right) \le \dots$$

Proof. The proof essentially follows the argument in the main result of Hajek's work [**?**], [**?**], [**?**] using Lemma .7 to, in the notation of Hajek, secure condition D1 and the fact that $\mathbb{E}\left[e^{\lambda\Delta\|u_t\|_2^2}\middle|\|u_{t-1}\|_2^2\right] \le e^{\frac{\lambda}{4}}$ in place of condition D2. For the sake of completeness, we include that argument here. Using the familiar Laplace transform trick with Chebyshev's inequality, we have for any $\lambda > 0$

$$\mathbb{P}\left(\|u_t\|_2^2 \ge \alpha\right) \le e^{-\lambda\alpha}\mathbb{E}\left[e^{\|u_t\|_2^2}\right] \le e^{-\lambda\alpha}\mathbb{E}\left[e^{\lambda\|u_{t-1}\|_2^2}e^{\lambda\Delta\|u_t\|_2^2}\right] = e^{-\lambda\alpha}\mathbb{E}\left[\mathbb{E}\left[e^{\lambda\|u_{t-1}\|_2^2}e^{\lambda\Delta\|u_t\|_2^2}\middle|\mathscr{F}_{t-1}\right]\right].$$

We can now slip this conditional expectation of the moment generating function of the increments into two pieces based on the norm of the previous residual as

$$\mathbb{E}\left[e^{\lambda\|u_{t-1}\|_2^2}e^{\lambda\Delta\|u_t\|_2^2}\middle|\mathscr{F}_{t-1}\right] = \mathbb{E}\left[e^{\lambda\|u_{t-1}\|_2^2}e^{\lambda\Delta\|u_t\|_2^2}\middle|\mathscr{F}_{t-1}, \|u_{t-1}\|_2 \ge \beta\right] + \mathbb{E}\left[e^{\lambda\|u_{t-1}\|_2^2}e^{\lambda\Delta\|u_t\|_2^2}\middle|\mathscr{F}_{t-1}, \|u_{t-1}\|_2 < \beta\right]$$

Specifically, we note by Lemma .7 when [****EL Note:** $\beta \gtrsim \sqrt{d}$] we have

$$\mathbb{E}\left[e^{\lambda\Delta\|u_t\|_2^2}\middle|\mathscr{F}_{t-1}, \|u_{t-1}\|_2^2 \ge \beta\right] = [\textbf{**EL Note:}\ \ \textit{some quantity...}] := \rho < 1. \tag{2}$$

Moreover, by Lemma .2 and Corollary .3

$$\mathbb{E}\left[e^{\lambda\|u_{t-1}\|_2^2}e^{\lambda\Delta\|u_t\|_2^2}\middle|\mathscr{F}_{t-1}, \|u_{t-1}\|_2^2 < \beta\right] \le e^{\frac{\lambda}{4}}e^{\lambda\beta}. \tag{3}$$

Combining (2), (3) then gives

$$\mathbb{P}\left(\|u_t\|_2^2 \ge \alpha\right) \le e^{-\lambda\alpha}\left(\rho\mathbb{E}\left[e^{\lambda\|u_{t-1}\|_2^2}\right] + e^{\frac{\lambda}{4}}\right).$$

Proceeding inductively thus yields

$$\mathbb{P}\left(\|u_t\|_2^2 \ge \alpha\right) \le \rho^t e^{\lambda\left(\|u_0\|_2^2 - \alpha\right)} + \frac{1-\rho^t}{1-\rho}e^{\lambda\left(\beta + \frac{1}{4} - \alpha\right)}.$$

[****EL Note:** *Go through and choose β, λ appropriately.*] □