APPENDIX

PROOFS

Lemma 2: Let A and B be measurable sets. If every measurable subset of A with non-zero measure is a subset of B, and if every measurable subset of B with non-zero measure is a subset of A, then $A = B$ almost everywhere.

Proof: We will prove this by contradiction. Let's assume that the hypothesis of the lemma holds, but $\mathrm{vol}(A \setminus B) > 0$. Since $A \setminus B = A \cap B^c$, it is a measurable set and $A \setminus B \subseteq A$, i.e., it is a measurable subset of A. By the hypothesis of the lemma, we have that $A \setminus B \subseteq B$ (as every measurable subset of A with non-zero measure is a subset of B). However, by definition, $A \setminus B$ is disjoint from B, therefore $A \setminus B = \emptyset \implies \mathrm{vol}(A \setminus B) = 0$, leading to a contradiction to our initial claim that $\mathrm{vol}(A \setminus B) > 0$. Therefore our initial claim must be wrong, and $\mathrm{vol}(A \setminus B) = 0$. Using the same argument $\mathrm{vol}(B \setminus A) = 0$, implying that $A = B$ almost everywhere. ∎

Proof: [**Proof of Theorem 1**] We establish this proof by showing that any measurable subset of F_t with non-zero measure is also a subset of ω_t^*, and every measurable subset of ω_t^* with non-zero measure is a subset of F_t and then invoking Lemma 2.

Let $\omega \subseteq \omega_t^*$ be a measurable set such that $\mathrm{vol}(\omega) > 0$. It follows that $\mu_t(\omega) > 0$ for if it doesn't, then $\mu_t(\omega_t^* \setminus \omega) = 1$, but $\mathrm{vol}(\omega_t^* \setminus \omega) < \mathrm{vol}(\omega_t^*)$ leading to a contradiction with the definition of ω_t^* in (3). Since $\mu_t(\omega) > 0$, by the properties of push-forward measure, there exist subsets $\mathcal{U}_i$ and $\mathcal{W}_i$ with non-zero measure that can bring the agent's state to ω. Therefore, $\omega \subseteq F_t$.

Let $F \subseteq F_t$ be a measurable set with $\mathrm{vol}(F) > 0$. As the composition of f in (2) for a fixed x_0—for the sake of convenience, let's call it $\hat{f}$—is a continuously differentiable map, by the Luzin's N-property for multi-dimensional maps, the pre-image $\hat{f}^{-1}(F) \subseteq \mathcal{U}_{t-1} \times \cdots \mathcal{U}_0 \times \mathcal{W}_{t-1} \times \cdots \times \mathcal{W}_0$ will be a set of non-zero Lebesgue measure as well. As we assume that the distribution over the space of controls and disturbances is absolutely continuous, and $\hat{f}^{-1}(F)$ lies in the support of these distributions with non-zero Lebesgue measure, it follows that the probability measure of $\hat{f}^{-1}(F)$ is also non-zero. By the property of the push-forward measure if the probability measure of $\hat{f}^{-1}(F)$ is non-zero, then the measure under the push-forward of these sets is also non-zero, i.e., $\mu_t(F) > 0$. Hence, $F \subseteq \omega_t^*$, for if it were removed from the set, the feasibility constraint of the lower bound will be violated. ∎

Proof: [**Proof of Theorem 2**] The volume of an ellipse $E_i(c_i)$ in an n-dimensional space [40] is given by

$$\mathrm{vol}(E_i(c_i)) = \frac{\pi^{n/2} c_i^{n/2}}{\Gamma(n/2+1)} \sqrt{\lambda_{i,1} \cdots \lambda_{i,n}} \tag{14}$$

where Γ is the standard gamma function in calculus while $\lambda_{i,j}$ are the eigen values of Σ. Setting $n = 2$ in the above and using it in (5) gives the objective function of (6).

In an n-dimensional space the function $V_i(x)$ is χ^2 distributed if x is drawn from a Gaussian [41]. For the special case of $n = 2$, the cumulative distribution of χ^2 gives $\mu_{t,i}(E_i(c_i)) = (1 - \exp(-c_i/2))$. Using this in (5) gives the constraint of (6).

Finally, we note that the objective function of (6) is linear, while the constraint is convex. For the latter assertion, note that the second derivative of $p_i(1 - \exp(-c_i/2))$ is $-p_i \exp(-c_i/2)/4$ which is always negative, hence $p_i(1 - \exp(-c_i/2))$ is concave for all i. Sum of concave functions is concave, therefore $\sum_{i=1}^K p_i(1 - \exp(-c_i/2))$ is concave. Since the super-level sets of concave functions are convex, we get that the constraint of (6) is also convex, implying that (6) is a convex optimization problem. ∎

Proof: [**Proof of Corollary 1**] We will establish this corollary by contradiction. Let $\alpha_1, \alpha_2 \in (0, \infty)$ and $\alpha_1 \neq \alpha_2$. Let $\{c_{i,\alpha_1}^*\}_{i=1}^K$ and $\{c_{i,\alpha_2}^*\}_{i=1}^K$ be the solutions of (6) for scaled covariances $\{\alpha_1 \Sigma_i\}_{i=1}^K$ and $\{\alpha_2 \Sigma_i\}_{i=1}^K$, respectively. Let's assume that there exists an i, such that $c_{i,\alpha_1}^* \neq c_{i,\alpha_2}^*$, and without loss of generality assume that $c_{i,\alpha_1}^* < c_{i,\alpha_2}^*$. Then,

$$\sum_{i=1}^K \pi \alpha_2 \sqrt{\lambda_{i,1}\lambda_{i,2}} c_{i,\alpha_1}^* < \sum_{i=1}^K \pi \alpha_2 \sqrt{\lambda_{i,1}\lambda_{i,2}} c_{i,\alpha_2}^*.$$

Further, observe that $\sum_{i=1}^K p_i(1 - \exp(-c_{i,\alpha_1}^*/2)) \geq \tau$ as it is an optimal (hence, feasible) solution of (6) for $\{\alpha_1 \Sigma_i\}_{i=1}^K$, therefore, $c_{1,\alpha_1}^*, \cdots, c_{K,\alpha_1}^*$ is also a feasible solution of (6) for $\{\alpha_1 \Sigma_i\}_{i=1}^K$, but attains a smaller objective function value than the optimal solution of (6) for α_2, leading to a contradiction, hence, the claim must be true. ∎

Proof: [**Proof of Lemma 1**] Using Corollary 1, it follows that $\{c_i^*\}_{i=1}^K$ are the same regardless of the scaling choice. As $\alpha_1 < \alpha_2$, then it follows that $\{x \mid (x - \overline{x}_i)^{\mathrm{T}} \alpha_1^{-1} \Sigma_i^{-1} (x - \overline{x}_i) \leq c_i^*\} \subseteq \{x \mid (x - \overline{x}_i)^{\mathrm{T}} \alpha_2^{-1} \Sigma_i^{-1} (x - \overline{x}_i) \leq c_i^*\}$ as $(x - \overline{x}_i)^{\mathrm{T}} \alpha_1^{-1} \Sigma_i^{-1} (x - \overline{x}_i) > (x - \overline{x}_i)^{\mathrm{T}} \alpha_2^{-1} \Sigma_i^{-1} (x - \overline{x}_i)$. Therefore, taking the union of all these sets gives us $E^*(\{\alpha_1 \Sigma_i\}_{i=1}^K) \subseteq E^*(\{\alpha_2 \Sigma_i\}_{i=1}^K)$. ∎