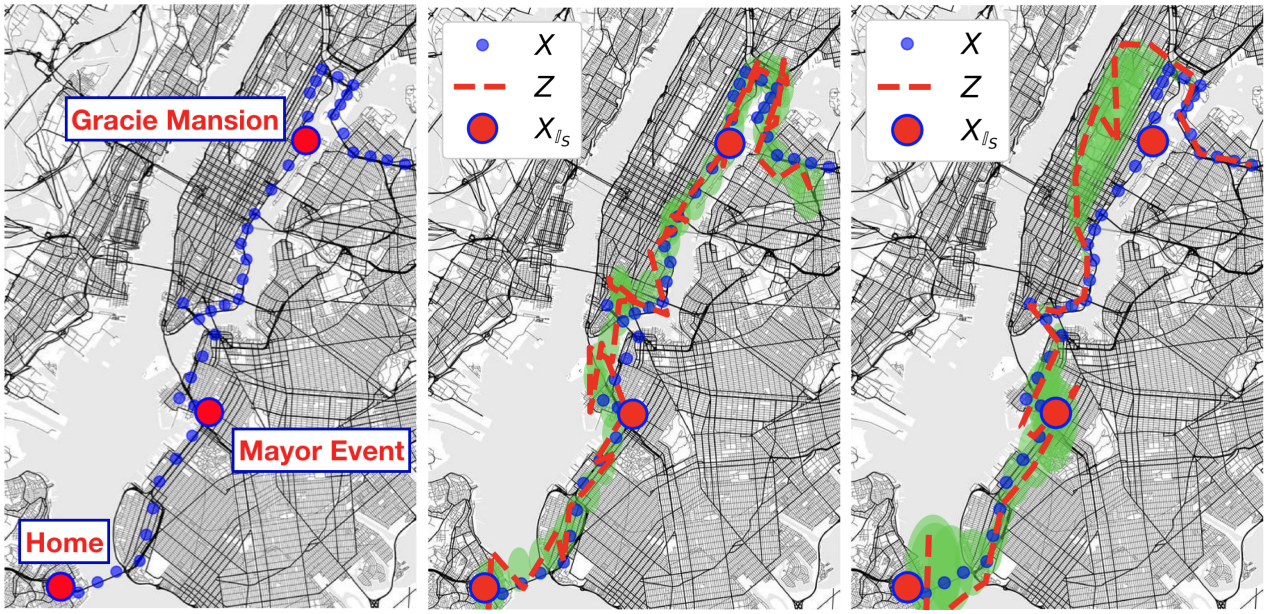
7 Appendix

For documented code demonstrating our SDP mechanisms used to generate the plots of **Figure 2** please visit our repo: `https://github.com/casey-meehan/location_trace_privacy`

The following sections will include proofs of results, derivations of algorithms, and explanations of experimental procedures.

7.1 Illustrations

7.1.1 NYC Mayoral Staff Member Location Trace



(a) **(b)** **(c)**

Figure 3: Example of sensitive location trace of NYC mayoral staff member exposed by (Valentino-DeVryes, 2018). (b) and (c) depict the posterior uncertainty (green) $P_{\mathcal{A},\mathcal{P}}(X_i|Z)$ for each 2d location. (a) depicts three sensitive times (red with blue outline): Gracie Mansion (Mayor's home), an event on Staten Island that the mayor attended, and finally the staff member's home on long island. (b) provides an example of Approach C: adding independent Gaussian noise to each location (red dotted line). A GP posterior still maintains high confidence within a small radius along the trace, including at the sensitive times. (c) provides an example of the optimized noise of Multiple Secrets of identical aggregate MSE as (b). By focusing *correlated* noise around the three sensitive times, there is high uncertainty at sensitive times and high confidence elsewhere.

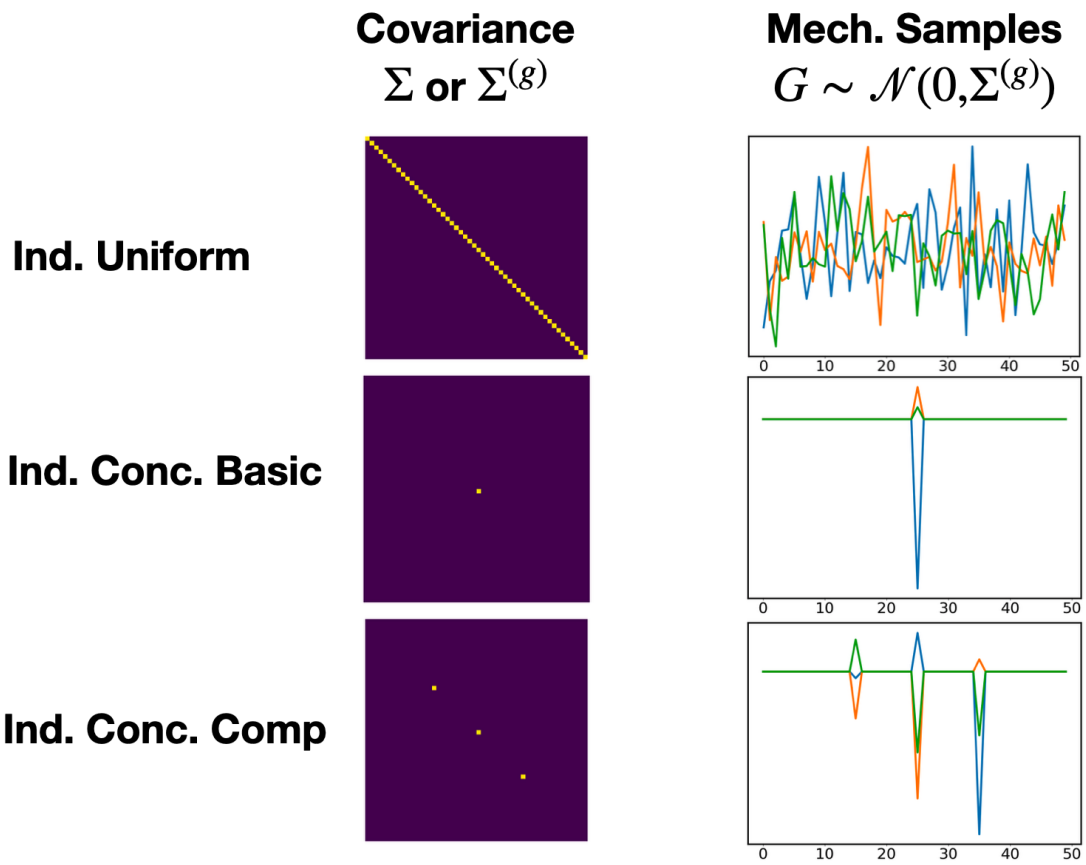
7.1.2 Juxtaposition of Mechanisms' Covariance Matrices

The following figures aim to illustrate the difference between the covariance matrices used in the experimental baselines (indep./uniform and indep./concentrated) and those chosen by our SDP algorithms for both the RBF and periodic prior. Note that here we presume the different dimensions of location to be independent and — by Corollary 7.2.1 — are able to treat a 2d location trace as two 1d traces. As such, the following examples are demonstrating mechanism covariance matrices and additive noise samples used for either a single dimension of location data (for RBF kernel) or for the one dimension of temperature data (for periodic kernel).

The first figure **(a)** shows the covariance of the Approach C baselines used in the experiments. The second figure **(b)** shows the covariance of our SDP mechanisms for the RBF kernel used on location data. The third figure **(c)** shows the covariance of our SDP mechanisms for the periodic kernel used for temperature data.

In each figure the covariance matrix is depicted as a heat map with warmer colors indicating higher values (normalized to largest and smallest value in the covariance matrix). The drawn noise samples G are plotted against their time index. So, the sequence of plotted (x, y) values is $\big[(1, G_1), (2, G_2), \ldots, (n, G_n)\big]$, where $n = 50$

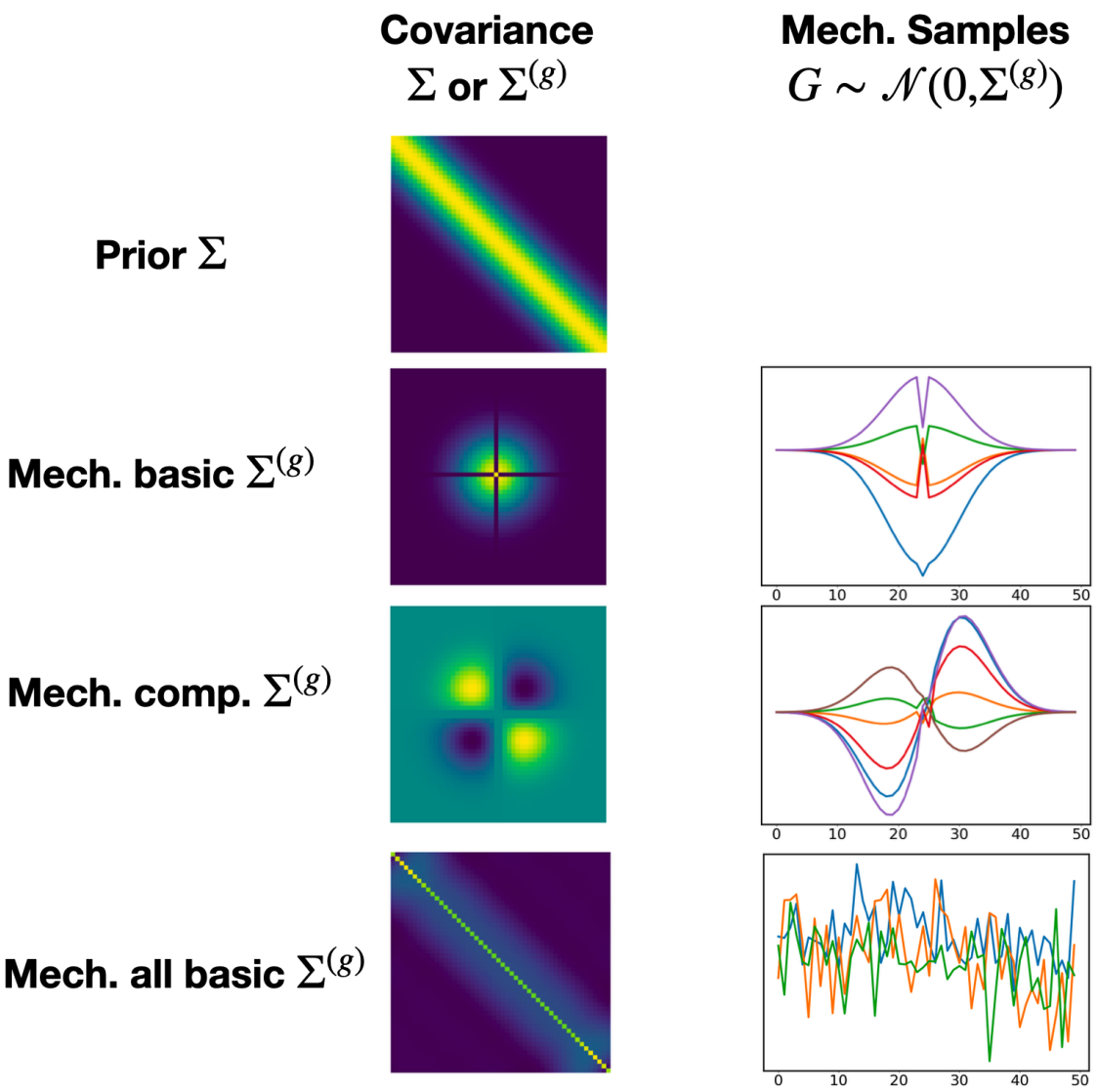
for the RBF case and $n = 48$ for the periodic case.



(a) Covariance matrices and mechanism samples for the baselines used in experiments.

The first figure demonstrates the uniform approach that distributes the independent Gaussian noise budget along the entire trace, regardless of $\mathbb{I}_S$.

The second and third show the concentrated approach that allocates the entire noise budget to only the sensitive locations in $\mathbb{I}_S$: first for a basic secret (one location) and then for a compound secret of 3 evenly spaced locations.

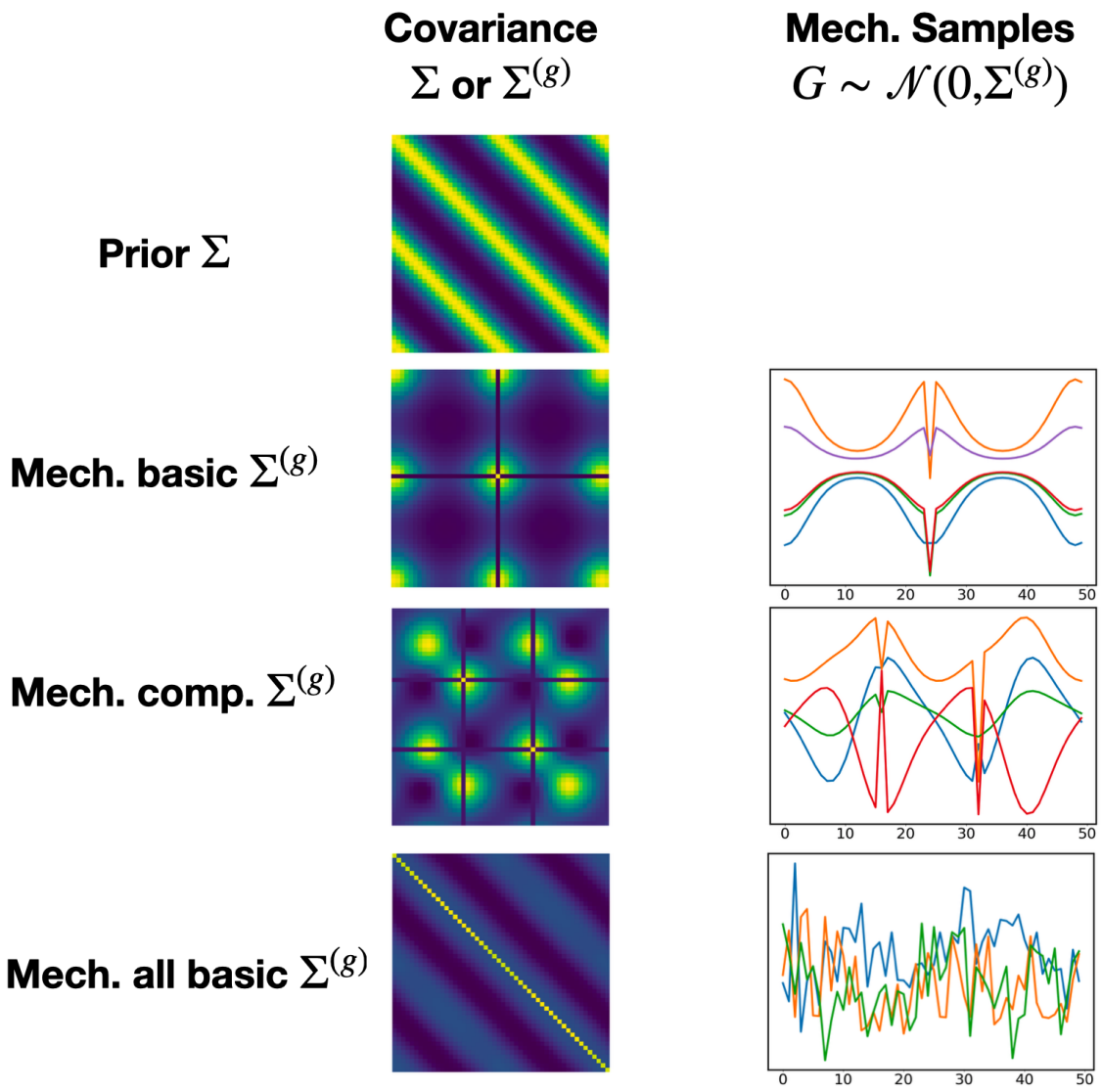


(b) Covariance matrices and mechanism samples for the median RBF prior ($l_{\text{eff}} \approx 6$).

The first noise mechanism (Mech. basic) demonstrates the covariance matrix chosen by SDP_A for a basic secret of a single location X_i in the middle of the trace. The uncorrelated dot in the middle of the covariance matrix, $\Sigma^{(g)}_{ii}$, represents the independent noise G_i added at the sensitive location to mitigate *direct* loss. To mitigate *inferential* loss, the SDP optimizes the remainder of the matrix to be positively correlated with maximum variance allocated to locations near X_i in time. This thwarts GP inference of the true location at time t_i.

The second mechanism (Mech. comp.) depicts the covariance chosen by SDP_A to protect a compound secret of two adjacent locations in the trace (visible as the uncorrelated '+' through the middle consuming 2 rows/columns). Recall that a compound secret ought to protect directional information: *did the user visit B first and then A, or A and then B?* That is precisely what this mechanism does by randomizing the angle of approach to the two locations in the middle with positively and negatively correlated noise. Also note that the SDP does not allocate a large share of noise budget to the actual locations themselves. This highlights the fact that protecting a compound secret does not protect its constituent basic secrets.

The third and final mechanism (Mech. all basic) is the noise covariance chosen by SDP_B in the Multiple Secrets algorithm. To protect all basic secrets with a utility constraint, the SDP converges to a mechanism that looks similar to the uniform baseline. However, this mechanism adds a subtle degree of off-diagonal correlation along with greater noise power towards the beginning and end of the trace. The off-diagonal correlation is noticeable when the samples are compared to those of the uniform baseline in the previous figure. While this change appears to be minor, it makes a significant change in the posterior confidence of a GP adversary (as seen in **Figure 2c**).



(c) Covariance matrices and mechanism samples for the median periodic prior ($l_{\text{eff}} \approx 1.1$), and a period of half the trace length.

The first noise mechanism (Mech. Basic) shows the covariance chosen by SDP_A to protect a single location (temperature) in the middle of the trace. As in the RBF case, significant noise power is allocated to the sensitive location itself, X_i, to limit *direct* privacy loss. However, the noise added to the remainder of the trace is significantly different. It is tailored to thwart inference by a periodic prior, wherein the location one period away has correlation 1.

The second noise mechanism (Mech. comp.) shows the covariance chosen by SDP_A to protect a compound secret of two locations, X_i, X_j, 16 timesteps apart (not quite a full period). Here, we see the SDP randomize the phase of the additive noise such that periodic inference cannot tell directional information like $X_i > X_j$ or vice versa.

The third noise mechanism (Mech. all basic) is identical to the all basic secrets mechanism chosen for the RBF case above, except using a periodic prior Σ. The mechanism chosen looks similar to the uniform baseline, except with slightly periodic off-diagonal correlation imitating the prior covariance. Additionally, noise power is mitigated towards the middle and ends of the trace. Again, **Figure 2g** indicates that this subtle change makes a significant difference in thwarting Bayesian adversaries.

7.2 Proof of results

7.2.1 Proof of Theorem 3.3

Theorem 3.3 Prior-Posterior Gap: *An (ε, λ)-CIP mechanism with conditional prior class Θ guarantees that for any event O on sanitized trace Z*

$$\left| \log \frac{P_{\mathcal{P},\mathcal{A}}(s_i | Z \in O)}{P_{\mathcal{P},\mathcal{A}}(s_j | Z \in O)} - \log \frac{P_{\mathcal{P}}(s_i)}{P_{\mathcal{P}}(s_j)} \right| \leq \varepsilon'$$

for any $\mathcal{P} \in \Theta$ with probability $\geq 1 - \delta$ over draws of $Z|X_{\mathbb{I}_S} = s_i$ or $Z|X_{\mathbb{I}_S} = s_j$, where ε' and δ are related by

$$\varepsilon' = \varepsilon + \frac{\log 1/\delta}{\lambda - 1} \ .$$

This holds under the condition that $Z|X_{\mathbb{I}_S} = s_i$ and $Z|X_{\mathbb{I}_S} = s_j$ have identical support.

Proof. This result makes use of a Rényi divergence property identified in Mironov (2017):

Lemma 7.1. *Let $\mathcal{P}, \mathcal{Q}$ be two distributions on X of identical support such that*

$$\max \left\{ D_\lambda \begin{pmatrix} P_{\mathcal{P}}(X) \\ P_{\mathcal{Q}}(X) \end{pmatrix}, D_\lambda \begin{pmatrix} P_{\mathcal{Q}}(X) \\ P_{\mathcal{P}}(X) \end{pmatrix} \right\} \leq \varepsilon$$

Then for any event O,

$$P_{\mathcal{P}}(X \in O) \leq \max \left\{ e^{\varepsilon'} P_{\mathcal{Q}}(X \in S), \delta \right\}$$

and

$$P_{\mathcal{Q}}(X \in O) \leq \max \left\{ e^{\varepsilon'} P_{\mathcal{P}}(X \in S), \delta \right\}$$

where

$$\varepsilon' = \varepsilon + \frac{\log 1/\delta}{\lambda - 1}$$

CIP guarantees that for all $\mathcal{P} \in \Theta$ and all discriminative pairs $(s_i, s_j) \in \mathcal{S}_{\text{pairs}}$ (which also includes (s_j, s_i))

$$D_\lambda \begin{pmatrix} P_{\mathcal{P},\mathcal{A}}(Z|X_{\mathbb{I}_S} = s_i) \\ P_{\mathcal{P},\mathcal{A}}(Z|X_{\mathbb{I}_S} = s_j) \end{pmatrix} \leq \varepsilon$$

and thus by Lemma 7.1 we have for any event O on Z

$$P_{\mathcal{P},\mathcal{A}}(Z \in O|X_{\mathbb{I}_S} = s_i) \leq \max \left\{ e^{\varepsilon'} P_{\mathcal{P},\mathcal{A}}(Z \in O|X_{\mathbb{I}_S} = s_j), \delta \right\}$$

and

$$P_{\mathcal{P},\mathcal{A}}(Z \in O|X_{\mathbb{I}_S} = s_j) \leq \max \left\{ e^{\varepsilon'} P_{\mathcal{P},\mathcal{A}}(Z \in O|X_{\mathbb{I}_S} = s_i), \delta \right\}$$

As such, given that $X_{\mathbb{I}_S} = s_i$ the probability of some event $\{Z \in W\}$ such that

$$P_{\mathcal{P},\mathcal{A}}(Z \in W|X_{\mathbb{I}_S} = s_i) \geq e^{\varepsilon'} P_{\mathcal{P},\mathcal{A}}(Z \in W|X_{\mathbb{I}_S} = s_j)$$

is no more than δ. The same is true swapping s_j for s_i. So, over draws of $Z|X_{\mathbb{I}_S} = s_i$ or $Z|X_{\mathbb{I}_S} = s_j$ we have that

$$\frac{P_{\mathcal{P},\mathcal{A}}(Z \in O|X_{\mathbb{I}_S} = s_i)}{P_{\mathcal{P},\mathcal{A}}(Z \in O|X_{\mathbb{I}_S} = s_j)} \leq e^{\varepsilon'} \quad \text{and} \quad \frac{P_{\mathcal{P},\mathcal{A}}(Z \in O|X_{\mathbb{I}_S} = s_j)}{P_{\mathcal{P},\mathcal{A}}(Z \in O|X_{\mathbb{I}_S} = s_i)} \leq e^{\varepsilon'}$$

with probability $\geq 1 - \delta$, which is equivalent to the statement that

$$-\varepsilon' \leq \log \frac{P_{\mathcal{P},\mathcal{A}}(Z \in O|X_{\mathbb{I}_S} = s_i)}{P_{\mathcal{P},\mathcal{A}}(Z \in O|X_{\mathbb{I}_S} = s_j)} \leq \varepsilon'$$

$$\left| \log \frac{P_{\mathcal{P},\mathcal{A}}(s_i|Z \in O)}{P_{\mathcal{P},\mathcal{A}}(s_j|Z \in O)} - \log \frac{P_{\mathcal{P}}(s_i)}{P_{\mathcal{P}}(s_j)} \right| \leq \varepsilon'$$

□

7.2.2 Proof of Lemma 3.2

Lemma 3.2 (CIP loss for additive mechanisms) *For an additive noise mechanism, a fully dependent trace as in **Figure 1b**, and any prior $\mathcal{P}$ on X the CIP loss may be expressed as*

$$D_\lambda\begin{pmatrix}P_{\mathcal{A},\mathcal{P}}(Z|X_{\mathbb{I}_S}=s_i)\\P_{\mathcal{A},\mathcal{P}}(Z|X_{\mathbb{I}_S}=s_j)\end{pmatrix} = \sum_{i\in\mathbb{I}_S}\left[D_\lambda\begin{pmatrix}P_{\mathcal{A}}(Z_i|X_i=s_i)\\P_{\mathcal{A}}(Z_i|X_i=s_j)\end{pmatrix}\right] + D_\lambda\begin{pmatrix}P_{\mathcal{A},\mathcal{P}}(Z_{\mathbb{I}_U}|X_{\mathbb{I}_S}=s_i)\\P_{\mathcal{A},\mathcal{P}}(Z_{\mathbb{I}_U}|X_{\mathbb{I}_S}=s_j)\end{pmatrix}$$

Proof.

$$\begin{aligned}
D_\lambda\begin{pmatrix}P_{\mathcal{A},\mathcal{P}}(Z|X_{\mathbb{I}_S}=x_s)\\P_{\mathcal{A},\mathcal{P}}(Z|X_{\mathbb{I}_S}=x'_s)\end{pmatrix} &= D_\lambda\begin{pmatrix}P_{\mathcal{A}}(Z_{\mathbb{I}_S}|X_{\mathbb{I}_S}=x_s)P_{\mathcal{A},\mathcal{P}}(Z_{\mathbb{I}_U}|X_{\mathbb{I}_S}=x_s)\\P_{\mathcal{A}}(Z_{\mathbb{I}_S}|X_{\mathbb{I}_S}=x'_s)P_{\mathcal{A},\mathcal{P}}(Z_{\mathbb{I}_U}|X_{\mathbb{I}_S}=x'_s)\end{pmatrix} &(1)\\
&= D_\lambda\begin{pmatrix}P_{\mathcal{A}}(Z_{\mathbb{I}_S}|X_{\mathbb{I}_S}=x_s)\\P_{\mathcal{A}}(Z_{\mathbb{I}_S}|X_{\mathbb{I}_S}=x'_s)\end{pmatrix} + D_\lambda\begin{pmatrix}P_{\mathcal{A},\mathcal{P}}(Z_{\mathbb{I}_U}|X_{\mathbb{I}_S}=x_s)\\P_{\mathcal{A},\mathcal{P}}(Z_{\mathbb{I}_U}|X_{\mathbb{I}_S}=x'_s)\end{pmatrix} &(2)\\
&= D_\lambda\begin{pmatrix}\prod_{i\in\mathbb{I}_S}P_{\mathcal{A}}(Z_i|X_i=x_i)\\\prod_{i\in\mathbb{I}_S}P_{\mathcal{A}}(Z_i|X_i=x'_i)\end{pmatrix} + D_\lambda\begin{pmatrix}P_{\mathcal{A},\mathcal{P}}(Z_{\mathbb{I}_U}|X_{\mathbb{I}_S}=x_s)\\P_{\mathcal{A},\mathcal{P}}(Z_{\mathbb{I}_U}|X_{\mathbb{I}_S}=x'_s)\end{pmatrix} &(3)\\
&= \sum_{i\in\mathbb{I}_S}\left[D_\lambda\begin{pmatrix}P_{\mathcal{A}}(Z_i|X_i=x_i)\\P_{\mathcal{A}}(Z_i|X_i=x'_i)\end{pmatrix}\right] + D_\lambda\begin{pmatrix}P_{\mathcal{A},\mathcal{P}}(Z_{\mathbb{I}_U}|X_{\mathbb{I}_S}=x_s)\\P_{\mathcal{A},\mathcal{P}}(Z_{\mathbb{I}_U}|X_{\mathbb{I}_S}=x'_s)\end{pmatrix} &(4)
\end{aligned}$$

Where line (1) uses the conditional independence seen in the graphical model of **Figure 1**. Line (2) is due to the fact that the two terms in line (1) are conditionally independent, allowing for separating into the sum of two separate divergences (which is an easily verifiable property of Rényi divergence evident from its definition in Equation 1). Line (3) is again from the conditional independence between the Z_i for each $i \in \mathbb{I}_S$ when conditioned on $X_{\mathbb{I}_S}$. Line (4) uses the same property of Rényi divergence used in Line (2): the terms in the product are conditionally independent allowing for the separation into the sum of multiple divergences.

□

7.2.3 Proof of Theorem 3.3

Thoerem 3.3 Robustness to Prior Misspecification *Mechanism $\mathcal{A}$ satisfies $\varepsilon(\lambda)$-CIP for prior class Θ. Suppose the finite mean true distribution $\mathcal{Q}$ is not in Θ. The CIP loss of $\mathcal{A}$ against prior $\mathcal{Q}$ is bounded by*

$$D_\lambda\begin{pmatrix}P_{\mathcal{A},\mathcal{Q}}(Z|X_{\mathbb{I}_S}=s_i)\\P_{\mathcal{A},\mathcal{Q}}(Z|X_{\mathbb{I}_S}=s_j)\end{pmatrix} \le \varepsilon'(\lambda)$$

where

$$\varepsilon'(\lambda) = \frac{\lambda-\frac{1}{2}}{\lambda-1}\,\Delta(2\lambda) + \Delta(4\lambda-3) + \frac{2\lambda-\frac{3}{2}}{2\lambda-2}\,\varepsilon(4\lambda-2)$$

and where $\Delta(\lambda)$ is

$$\inf_{\mathcal{P}\in\Theta}\sup_{s_i\in\mathcal{S}}\max\left\{D_\lambda\begin{pmatrix}P_{\mathcal{P}}(X_{\mathbb{I}_U}|X_{\mathbb{I}_S}=s_i)\\P_{\mathcal{Q}}(X_{\mathbb{I}_U}|X_{\mathbb{I}_S}=s_i)\end{pmatrix}, D_\lambda\begin{pmatrix}P_{\mathcal{Q}}(X_{\mathbb{I}_U}|X_{\mathbb{I}_S}=s_i)\\P_{\mathcal{P}}(X_{\mathbb{I}_U}|X_{\mathbb{I}_S}=s_i)\end{pmatrix}\right\}$$

Proof. By 'finite mean' distribution $\mathcal{Q}$, we mean that all conditionals of $\mathcal{Q}$ given some $X_{\mathbb{I}_S}$ have finite mean. Since a conditional prior class contains conditionals of one distribution with any offset (any mean value), this guarantees that $\Delta(\lambda)$ is achieved for some $\mathcal{P} \in \Theta$. Intuitively, this prevents the pathological case of $\inf_{\mathcal{P}\in\Theta}$ being a limit as the mean of $\mathcal{P} \to \infty$, only asymptotically approaching $\Delta(\lambda)$. If the mean of $\mathcal{Q}$ is finite, then the closest $\mathcal{P} \in \Theta$ (in Rényi divergence) must also have finite mean, since any mean is attainable in a conditional prior class Θ.

With this in mind, we make use of the following triangle inequality provided in Mironov (2017):

Lemma 7.2. *For distributions $\mathcal{P}$, $\mathcal{Q}$, $\mathcal{R}$ on X with common support we have*

$$D_\lambda\binom{P_\mathcal{P}(X)}{P_\mathcal{Q}(X)} \leq \frac{\lambda-\frac{1}{2}}{\lambda-1} D_{2\lambda}\binom{P_\mathcal{P}(X)}{P_\mathcal{R}(X)} + D_{2\lambda-1}\binom{P_\mathcal{R}(X)}{P_\mathcal{Q}(X)}$$

In our case, we assume that the mechanism $\mathcal{A}$ gives $Z|X_{\mathbb{I}_S} = x_s$ identical support for all $\mathbb{I}_S, x_s$. Using this, we have

$$D_\lambda\binom{P_{\mathcal{A},\mathcal{Q}}(Z_{\mathbb{I}_U}|X_{\mathbb{I}_S}=x_s)}{P_{\mathcal{A},\mathcal{Q}}(Z_{\mathbb{I}_U}|X_{\mathbb{I}_S}=x'_s)} \leq \frac{\lambda-\frac{1}{2}}{\lambda-1} D_{2\lambda}\binom{P_{\mathcal{A},\mathcal{Q}}(Z_{\mathbb{I}_U}|X_{\mathbb{I}_S}=x_s)}{P_{\mathcal{A},\mathcal{P}}(Z_{\mathbb{I}_U}|X_{\mathbb{I}_S}=x_s)} + D_{2\lambda-1}\binom{P_{\mathcal{A},\mathcal{P}}(Z_{\mathbb{I}_U}|X_{\mathbb{I}_S}=x_s)}{P_{\mathcal{A},\mathcal{Q}}(Z_{\mathbb{I}_U}|X_{\mathbb{I}_S}=x'_s)} \quad .$$

By a data processing inequality, the divergence of the first term is bounded by $\Delta(2\lambda)$ and the blue term may be bounded by a second application of the triangle inequality:

$$D_{2\lambda-1}\binom{P_{\mathcal{A},\mathcal{P}}(Z_{\mathbb{I}_U}|X_{\mathbb{I}_S}=x_s)}{P_{\mathcal{A},\mathcal{Q}}(Z_{\mathbb{I}_U}|X_{\mathbb{I}_S}=x'_s)} \leq \frac{2\lambda-\frac{3}{2}}{2\lambda-2} D_{4\lambda-2}\binom{P_{\mathcal{A},\mathcal{P}}(Z_{\mathbb{I}_U}|X_{\mathbb{I}_S}=x_s)}{P_{\mathcal{A},\mathcal{P}}(Z_{\mathbb{I}_U}|X_{\mathbb{I}_S}=x'_s)} + D_{4\lambda-3}\binom{P_{\mathcal{A},\mathcal{P}}(Z_{\mathbb{I}_U}|X_{\mathbb{I}_S}=x'_s)}{P_{\mathcal{A},\mathcal{Q}}(Z_{\mathbb{I}_U}|X_{\mathbb{I}_S}=x'_s)}$$

The first divergence is bounded by $\varepsilon(4\lambda-2)$ and the second divergence is bounded by $\Delta(4\lambda-3)$. Putting all this together we have the following upper bound

$$D_\lambda\binom{P_{\mathcal{A},\mathcal{Q}}(Z_{\mathbb{I}_U}|X_{\mathbb{I}_S}=x_s)}{P_{\mathcal{A},\mathcal{Q}}(Z_{\mathbb{I}_U}|X_{\mathbb{I}_S}=x'_s)} \leq \frac{\lambda-\frac{1}{2}}{\lambda-1}\,\Delta(2\lambda) + \Delta(4\lambda-3) + \frac{2\lambda-\frac{3}{2}}{2\lambda-2}\,\varepsilon(4\lambda-2)$$

□

7.2.4 Proof of Theorem 3.4

Theorem 3.4 CIP loss bound for GP conditional priors: *Let Θ be a GP conditional prior class. Let Σ be the covariance matrix for X produced by its kernel function. Let $\mathcal{S}$ be the basic or compound secret associated with $\mathbb{I}_S$, and S be the number of unique times in $\mathbb{I}_S$. The mechanism $\mathcal{A}(X) = X + G = Z$, where $G \sim \mathcal{N}(\mathbf{0}, \Sigma^{(g)})$, then satisfies (ε, λ)-Conditional Inferential Privacy $(\mathcal{S}_{pairs}, r, \Theta)$, where*

$$\varepsilon \leq \frac{\lambda}{2} S r^2 \Big(\frac{1}{\sigma_s^2} + \alpha^*\Big)$$

where σ_s^2 is the variance of each $G_i \in G_{\mathbb{I}_S}$ (diagonal entries of $\Sigma_{ss}^{(g)}$) and α^* is the maximum eigenvalue of $\Sigma_{eff} = \big(\Sigma_{us}\Sigma_{ss}^{-1}\big)^\intercal \big(\Sigma_{u|s} + \Sigma_{uu}^{(g)}\big)^{-1}\big(\Sigma_{us}\Sigma_{ss}^{-1}\big)$.

Proof. Again, the conditional prior class Θ is defined by a kernel function $i, j \to \mathrm{Cov}(i, j)$, which – given the indices of the trace X – induces a covariance matrix Σ between all X_i, X_j. In practice, when the sampling rate of locations is non-uniform the kernel function may use the time-stamps of the points in the trace to assign high correlation to X_i that are close in time and low correlation to X_i that are far apart in time. Of course, correlation between X_i that are different dimension (e.g. latitude and longitude) must be designed for the given application and may be completely independent. The kernel function can encode this as well.

Recall from Equation 1 that the Rényi divergence between two mean-shifted multivariate normal distributions, $\mathcal{P}_1 = \mathcal{N}(\mu_1, \Sigma)$ and $\mathcal{P}_2 = \mathcal{N}(\mu_2, \Sigma)$ is

$$D_\lambda\binom{\mathcal{P}_1}{\mathcal{P}_2} = \frac{\lambda}{2}(\mu_1 - \mu_2)^\intercal \Sigma^{-1}(\mu_1 - \mu_2)$$

Now, for any prior $\mathcal{P} \in \Theta$, we have that $X \sim \mathcal{N}(\mu, \Sigma)$ for some μ and for Σ defined by the kernel function. Again, $G \sim \mathcal{N}(\mathbf{0}, \Sigma^{(g)})$. $\mathbb{I}_S$ encodes the indices of a single location basic secret or a multi-location compound secret. Then, the divergence to bound for (ε, λ)-CIP$(\mathcal{S}_{\text{pairs}}, r, \Theta)$ is

$$D_\lambda\binom{P_{\mathcal{A},\mathcal{P}}(Z|X_{\mathbb{I}_S}=s_i)}{P_{\mathcal{A},\mathcal{P}}(Z|X_{\mathbb{I}_S}=s_j)}$$

for any

$$(s_i, s_j) \in \mathcal{S}_{\text{pairs}} = \{(x_s, x'_s) : \|x_s - x'_s\|_2 \leq 2r\}$$

if $\mathbb{I}_S$ encodes a basic secret, or for any

$$(s_i, s_j) \in \mathcal{S}_{\text{pairs}} = \left\{ \left(\{x_{s1}, x_{s2}, \dots\}, \{x'_{s1}, x'_{s2}, \dots\}\right) : \|x_{sk} - x'_{sk}\|_2 \leq 2r, \forall\ k \right\}$$

if $\mathbb{I}_S$ encodes a compound secret. A discriminative pair (s_i, s_j) is two real valued vectors $\in \mathbb{R}^{|\mathbb{I}_S|}$, representing two hypotheses about the true values of $X_{\mathbb{I}_S}$. We denote the m^{th} element as $s_{im}, s_{j\,m}$. Let $f : \mathbb{I}_S \to [|\mathbb{I}_S|]$ be a mapping from each index $w \in \mathbb{I}_S$ to its corresponding position in the vector s_i or s_j (where the value of X_w is hypothesized). By Lemma 3.2, the divergence can be written as

$$D_\lambda\begin{pmatrix} P_{\mathcal{A},\mathcal{P}}(Z|X_{\mathbb{I}_S} = s_i) \\ P_{\mathcal{A},\mathcal{P}}(Z|X_{\mathbb{I}_S} = s_j) \end{pmatrix} = \sum_{w \in \mathbb{I}_S} \left[D_\lambda\begin{pmatrix} P_{\mathcal{A}}(Z_w|X_w = s_{i\,f(w)}) \\ P_{\mathcal{A}}(Z_w|X_w = s_{j\,f(w)}) \end{pmatrix} \right] + D_\lambda\begin{pmatrix} P_{\mathcal{A},\mathcal{P}}(Z_{\mathbb{I}_U}|X_{\mathbb{I}_S} = x_s) \\ P_{\mathcal{A},\mathcal{P}}(Z_{\mathbb{I}_U}|X_{\mathbb{I}_S} = x'_s) \end{pmatrix}$$

where $P_{\mathcal{A}}(Z_w|X_w = x) = \mathcal{N}(x, \sigma_s^2)$ for all $w \in \mathbb{I}_S$. Recall from the statement of the Theorem that we assume the diagonal entries of Σ_{ss} all equal some value σ_s^2: we add the same noise variance to each point in the secret set, which is optimal under MSE constraints. Additionally, note that for the hypothesis $X_{\mathbb{I}_S} = x_s$, we know the distribution of $X_{\mathbb{I}_U}|X_{\mathbb{I}_S} = x_s \sim \mathcal{N}(\mu_{u|s}, \Sigma_{u|s})$, where $\mu_{u|s} = \mu_u + \Sigma_{us}\Sigma_{ss}^{-1}(x_s - \mu_s)$ and $\Sigma_{u|s} = \Sigma_{uu} - \Sigma_{us}\Sigma_{ss}^{-1}\Sigma_{su}$. Notice that only $\mu_{u|s}$ depends on the actual value of x_s, and $\Sigma_{u|s}$ depends only on the indices of $\mathbb{I}_S$. Being the sum of two normally distributed variables, we have that $(Z_{\mathbb{I}_U}|X_{\mathbb{I}_S} = x_s) \overset{d}{=} (X_{\mathbb{I}_U}|X_{\mathbb{I}_S} = x_s) + G_{\mathbb{I}_U} = \mathcal{N}(\mu_{u|s}, \Sigma_{u|s} + \Sigma_{uu}^{(g)})$. Substituting this into the divergences above sum of divergences:

$$D_\lambda\begin{pmatrix} P_{\mathcal{A},\mathcal{P}}(Z|X_{\mathbb{I}_S} = s_i) \\ P_{\mathcal{A},\mathcal{P}}(Z|X_{\mathbb{I}_S} = s_j) \end{pmatrix} = \sum_{m=1}^{|\mathbb{I}_S|} \left[D_\lambda\begin{pmatrix} \mathcal{N}(s_{im}, \sigma_s^2) \\ \mathcal{N}(s_{j\,m}, \sigma_s^2) \end{pmatrix} \right] + D_\lambda\begin{pmatrix} \mathcal{N}(\mu_{u|s_i}, \Sigma_{u|s} + \Sigma_{uu}^{(g)}) \\ \mathcal{N}(\mu_{u|s_j}, \Sigma_{u|s} + \Sigma_{uu}^{(g)}) \end{pmatrix} \tag{1}$$

$$= \frac{\lambda}{2}\sum_{m=1}^{|\mathbb{I}_S|} \frac{1}{\sigma_s^2}(s_{im} - s_{j\,m})^2 + \frac{\lambda}{2}(\mu_{u|s_i} - \mu_{u|s_j})^\intercal(\Sigma_{u|s} + \Sigma_{uu}^{(g)})^{-1}(\mu_{u|s_i} - \mu_{u|s_j}) \tag{2}$$

$$= \frac{\lambda}{2\sigma_s^2}(s_i - s_j)^\intercal(s_i - s_j) + \frac{\lambda}{2}\big(\Sigma_{us}\Sigma_{ss}^{-1}(s_i - s_j)\big)^\intercal(\Sigma_{u|s} + \Sigma_{uu}^{(g)})^{-1}\big(\Sigma_{us}\Sigma_{ss}^{-1}(s_i - s_j)\big) \tag{3}$$

$$= \frac{\lambda}{2\sigma_s^2}(s_i - s_j)^\intercal(s_i - s_j) + \frac{\lambda}{2}(s_i - s_j)^\intercal\Sigma_{ss}^{-1}\Sigma_{su}(\Sigma_{u|s} + \Sigma_{uu}^{(g)})^{-1}\Sigma_{us}\Sigma_{ss}^{-1}(s_i - s_j) \tag{4}$$

Line (1) substitutes in the normal distributions given by our mechanism and conditional prior class. Line (2) substitutes in the closed-form expression for Rényi divergence between two mean-shifted normal distributions given in Equation 1. Line (3) substitutes in the expression for $\mu_{u|s}$ given above, and simplifies. To expand out this simplification in explicit steps:

$$\begin{aligned}(\mu_{u|s_i} - \mu_{u|s_j}) &= \big(\mu_u + \Sigma_{us}\Sigma_{ss}^{-1}(s_i - \mu_s) - [\mu_u + \Sigma_{us}\Sigma_{ss}^{-1}(s_j - \mu_s)]\big) \\ &= \big(\Sigma_{us}\Sigma_{ss}^{-1}s_i - \Sigma_{us}\Sigma_{ss}^{-1}s_j\big) \\ &= \Sigma_{us}\Sigma_{ss}^{-1}(s_i - s_j)\end{aligned}$$

Line (4) distributes the transpose in the right term of line (3):

$$\begin{aligned}\big(\Sigma_{us}\Sigma_{ss}^{-1}(s_i - s_j)\big)^\intercal &= (s_i - s_j)^\intercal\big(\Sigma_{us}\Sigma_{ss}^{-1}\big)^\intercal \\ &= (s_i - s_j)^\intercal\big(\Sigma_{ss}^{-1}\big)^\intercal\Sigma_{us}^\intercal \\ &= (s_i - s_j)^\intercal\Sigma_{ss}^{-1}\Sigma_{su}\end{aligned}$$

where that final step is a consequence of Σ being symmetric. Σ_{ss} is also a symmetric matrix (so its inverse is symmetric) and $\Sigma_{us}^\intercal = \Sigma_{su}$.

Returning to line (4) above, simplify this expression by substituting $\Delta = s_i - s_j$:

$$D_\lambda\begin{pmatrix} P_{\mathcal{A},\mathcal{P}}(Z|X_{\mathbb{I}_S} = s_i) \\ P_{\mathcal{A},\mathcal{P}}(Z|X_{\mathbb{I}_S} = s_j) \end{pmatrix} = \frac{\lambda}{2\sigma_s^2}\Delta^\intercal\Delta + \frac{\lambda}{2}\Delta^\intercal\Sigma_{ss}^{-1}\Sigma_{su}(\Sigma_{u|s} + \Sigma_{uu}^{(g)})^{-1}\Sigma_{us}\Sigma_{ss}^{-1}\Delta \tag{5}$$

$$= \frac{\lambda}{2\sigma_s^2}\|\Delta\|_2^2 + \frac{\lambda}{2}\Delta^\intercal\Sigma_{\text{eff}}\Delta \tag{6}$$

Where $\Sigma_{\text{eff}} = \Sigma_{ss}^{-1}\Sigma_{su}(\Sigma_{u|s} + \Sigma_{uu}^{(g)})^{-1}\Sigma_{us}\Sigma_{ss}^{-1}$. The left term of line (6) attributes the direct loss of $Z_{\mathbb{I}_S}$ on $X_{\mathbb{I}_S}$ and the right term attributes the indirect loss of $Z_{\mathbb{I}_U}$ on $X_{\mathbb{I}_S}$.

We are interested in bounding the expression of line (6) for all $(s_i, s_j) \in \mathcal{S}_{\text{pairs}}$. We do this by bounding it for all vectors $\Delta \in \mathcal{D}$

$$\mathcal{D} = \{s_i - s_j : \|s_i - s_j\|_2 \leq \sqrt{S}\ r\}$$

, where S is the number of basic secrets (locations) contained in $\mathbb{I}_S$ which may be a basic or compound secret set. For a basic secret ($S = 1$), this bound is tight, since $\mathcal{D} = \{s_i - s_j : (s_i, s_j) \in \mathcal{S}_{\text{pairs}}\}$. The set of $\Delta \in \mathcal{D}$ is exactly any two hypothesis (s_i, s_j) that are within any circle of radius r. For a compound secret, this bound is not guaranteed to be tight. Recall once again that the set of $\mathcal{S}_{\text{pairs}}$ for a compound secret is given by the set of (s_i, s_j) in

$$\mathcal{S}_{\text{pairs}} = \left\{ \left(\{x_{s1}, x_{s2}, \dots\}, \{x'_{s1}, x'_{s2}, \dots\} \right) : \|x_{sk} - x'_{sk}\|_2 \leq r, \forall\ k \right\}$$

For concreteness, consider the 2d location trace example in **Figure 3**, where we have a compound secret of $S = 3$ locations. Here, $s_i, s_j \in \mathbb{R}^6$, where 6 comes from the fact that we have three 2d locations. So, (s_i, s_j) represents a pair of hypotheses on all three locations. s_i's hypothesis of the first secret location — written as $x_{s1} \in \mathbb{R}^2$ above — is within r of the s_j's hypothesis of the first secret location — written as $x_{s_1'} \in \mathbb{R}^2$ above. The same goes for the second and third locations. So, the L_2 norm of $\Delta = s_i - s_j$ is no greater than

$$\begin{aligned}
\sup_{(s_i,s_j)\in\mathcal{S}_{\text{pairs}}} \|s_i - s_j\|_2 &= \sup_{(s_i,s_j)\in\mathcal{S}_{\text{pairs}}} \sqrt{\sum_{m=1}^{6} (s_{i_m} - s_{j_m})^2} \\
&= \sup_{(s_i,s_j)\in\mathcal{S}_{\text{pairs}}} \sqrt{\sum_{k=1}^{3} \|x_{sk} - x_{s_k'}\|_2^2} \\
&= \sqrt{\sum_{k=1}^{3} r^2} \\
&= \sqrt{3}\ r
\end{aligned}$$

For compound secrets, $\mathcal{D}$ represents the L_2 ball enclosing all $\Delta \in \{s_i - s_j : (s_i, s_j) \in \mathcal{S}_{\text{pairs}}\}$. However, $\mathcal{D}$ also includes some values of $\Delta = s_i - s_j$ not covered by $\mathcal{S}_{\text{pairs}}$. Suppose an adversary considers the hypotheses

$$s_i = \{x_{s1}, x_{s2}, x_{s3}\}, s_j = \{x'_{s1}, x'_{s2}, x'_{s3}\}$$

where $x_{s1} = 0, x_{s1}' = \sqrt{3}\ r, x_{s2} = x_{s2}', x_{s3} = x_{s3}'$. Since x_{s1}, x_{s1}' are not within r of each other, this is not in $\mathcal{S}_{\text{pairs}}$. However, it is covered by $\mathcal{D}$, and thus is covered by our bound on CIP loss and our mechanisms.

With $\mathcal{D}$ defined, we may return to bounding the expression in line (6):

$$\begin{aligned}
D_\lambda \begin{pmatrix} P_{\mathcal{A},\mathcal{P}}(Z|X_{\mathbb{I}_S} = s_i) \\ P_{\mathcal{A},\mathcal{P}}(Z|X_{\mathbb{I}_S} = s_j) \end{pmatrix} &\leq \sup_{\Delta\in\mathcal{D}} \left(\frac{\lambda}{2\sigma_s^2}\|\Delta\|_2^2 + \frac{\lambda}{2}\Delta^\intercal\Sigma_{\text{eff}}\Delta \right) && (7) \\
&\leq \frac{\lambda}{2}\left(\frac{1}{\sigma_s^2} S r^2 + S r^2 \text{maxeig}(\Sigma_{\text{eff}}) \right) && (8) \\
&= \frac{\lambda}{2} S r^2 (\frac{1}{\sigma_s^2} + \alpha^*) && (9)
\end{aligned}$$

where line (8) distributes the supremum. For the right term, this is given by the maximum magnitude of all $\Delta \in \mathcal{D}$ times the maximum eigenvalueof Σ_{eff} which equals $Sr^2\text{maxeig}(\Sigma_{\text{eff}})$. Line (9) simply substitutes $\alpha^* = \text{maxeig}(\Sigma_{\text{eff}})$.

□

7.2.5 Proof of Corollary 3.4.1

Corollary 3.4.1 Graceful Composition in Time *Suppose a user releases two traces X and $\hat{X}$ with additive noise $G \sim \mathcal{N}(\mathbf{0}, \Sigma^{(g)})$ and $\hat{G} \sim \mathcal{N}(\mathbf{0}, \hat{\Sigma}^{(g)})$, respectively. Then basic or compound secret $X_{\mathbb{I}_S}$ of X enjoys $(\bar{\varepsilon}, \lambda)$-CIP, where*

$$\bar{\varepsilon} \leq \frac{\lambda}{2} S r^2 \Big(\frac{1}{\sigma_s^2} + \bar{\alpha}^* \Big)$$

and where $\bar{\alpha}$ is the maximum eigenvalue of $\bar{\Sigma}_{\mathit{eff}} = \big(\Sigma_{us}\Sigma_{ss}^{-1}\big)^{\intercal}\big(\Sigma_{u|s} + \bar{\Sigma}_{uu}^{(g)}\big)^{-1}\big(\Sigma_{us}\Sigma_{ss}^{-1}\big)$. Σ is the covariance matrix of the joint distribution on $X, \hat{X}$ and

$$\bar{\Sigma}^{(g)} = \begin{bmatrix} \Sigma^{(g)} & 0 \\ 0 & \hat{\Sigma}^{(g)} \end{bmatrix}.$$

Proof. Here, we record two traces (presumably) far apart in time

$$(X_1, \ldots, X_n) \text{ and } (\hat{X}_1, \ldots, \hat{X}_m)$$

And release

$$(Z_1, \ldots, Z_n) = (X_1, +G_1, \ldots, X_n + G_n) \text{ and } (\hat{Z}_1, \ldots, \hat{Z}_m) = (\hat{X}_1, +\hat{G}_1, \ldots, \hat{X}_m, +\hat{G}_m)$$

the first trace protects secret locations $X_{\mathbb{I}_S}$ and the second protects $\widehat{X_{\mathbb{I}_S}}$, so we have that

$$D_\lambda \begin{pmatrix} P_{\mathcal{A},\mathcal{P}}(Z | X_{\mathbb{I}_S} = s_i) \\ P_{\mathcal{A},\mathcal{P}}(Z | X_{\mathbb{I}_S} = s_j) \end{pmatrix} \leq \varepsilon$$
$$D_\lambda \begin{pmatrix} P_{\mathcal{A},\mathcal{P}}(\hat{Z} | \widehat{X_{\mathbb{I}_S}} = \hat{s}_i) \\ P_{\mathcal{A},\mathcal{P}}(\hat{Z} | \widehat{X_{\mathbb{I}_S}} = \hat{s}_j) \end{pmatrix} \leq \hat{\varepsilon}$$

We aim to update the losses:

$$D_\lambda \begin{pmatrix} P_{\mathcal{A},\mathcal{P}}(Z, \hat{Z} | X_{\mathbb{I}_S} = s_i) \\ P_{\mathcal{A},\mathcal{P}}(Z, \hat{Z} | X_{\mathbb{I}_S} = s_j) \end{pmatrix} \leq \varepsilon'$$
$$D_\lambda \begin{pmatrix} P_{\mathcal{A},\mathcal{P}}(\hat{Z}, Z | \widehat{X_{\mathbb{I}_S}} = \hat{s}_i) \\ P_{\mathcal{A},\mathcal{P}}(\hat{Z}, Z | \widehat{X_{\mathbb{I}_S}} = \hat{s}_j) \end{pmatrix} \leq \hat{\varepsilon}'$$

Fortunately, our framework is pretty friendly to figuring this out, and can be done simply by updating the 'inferential loss term' α^* and $\hat{\alpha}^*$ of each, the max eigenvalues used to compute each of ε and $\hat{\varepsilon}$, respectively. Let's focus on ε', since the same analysis follows for $\hat{\varepsilon}'$.

Recall that α^* is given by the max eigenvalue of Σ_{eff} which is

$$\Sigma_{\text{eff}} = \big(\Sigma_{us}\Sigma_{ss}^{-1}\big)^{\intercal}\big(\Sigma_{u|s} + \Sigma_{uu}^{(g)}\big)^{-1}\big(\Sigma_{us}\Sigma_{ss}^{-1}\big)$$

Where Σ is the covariance matrix of $X_1, \ldots, X_n$ and $\Sigma^{(g)}$ is the noise covariance matrix added. Simply augment Σ to become the joint covariance matrix Σ_J of $X, \hat{X}$, and augment $\Sigma^{(g)}$ to become

$$\Sigma_J^{(g)} = \begin{bmatrix} \Sigma^{(g)} & 0 \\ 0 & \hat{\Sigma}^{(g)} \end{bmatrix}$$

then update Σ_{eff} to $\Sigma_{\text{eff},J}$ which uses both Σ_J and $\Sigma_J^{(g)}$. Using the corresponding max eigenvalue α_J^* in the loss expression of Theorem 3.2 gives us ε'.

Note that for kernels like RBF, $\varepsilon' \to \varepsilon$ as the traces X and $\hat{X}$ move apart further and further in time. This is not the case for traces using a purely periodic kernel with not time decay, and we should expect much worse composition. □

7.2.6 Traces with Independent Dimensions

In many cases, the different dimensions of the trace may be probabilistically independent, and it may be more convenient to make separate privacy mechanisms for each. For a 2d trace X, suppose $\mathbb{I}_x$ and $\mathbb{I}_y$ store the indices of the latitude points $X_{\mathbb{I}_x}$ and longitude points $X_{\mathbb{I}_y}$, such that $X = X_{\mathbb{I}_x} \cup X_{\mathbb{I}_y}$. If latitude and longitude are independent, it may be more convenient to characterize the conditional priors of $X_{\mathbb{I}_x}$ abd $X_{\mathbb{I}_y}$ separately. The question is whether privacy guarantees remain for the full trace X. To answer this, we provide the following corollary:

Corollary 7.2.1. CIP loss of independent dimensions *Let Θ be a GP conditional prior class on a 2d trace X such that the dimensions are independent. Let $\mathbb{I}_S$ be some secret set of time indices corresponding to some basic or compound secret. For the trace $X = X_{\mathbb{I}_x} \cup X_{\mathbb{I}_y}$, the Gaussian mechanism $\mathcal{A}(X) = Z_{\mathbb{I}_x} \cup Z_{\mathbb{I}_y}$ where $Z_{\mathbb{I}_x} = \mathcal{A}_x(X_{\mathbb{I}_x}) = X_{\mathbb{I}_x} + G_{\mathbb{I}_x}$ and $Z_{\mathbb{I}_y} = \mathcal{A}_y(X_{\mathbb{I}_y}) = X_{\mathbb{I}_y} + G_{\mathbb{I}_y}$ satisfies (ε, λ)-CIP where*

$$\varepsilon \leq \frac{\lambda}{2} Sr^2 \big(\frac{1}{\sigma_s^2} + \alpha_x^* + \alpha_y^* \big)$$

when $\mathcal{A}_x$ and $\mathcal{A}_y$ provide $\frac{\lambda}{2} Sr^2 \big(\frac{1}{\sigma_s^2} + \alpha_x^* \big)$ and $\frac{\lambda}{2} Sr^2 \big(\frac{1}{\sigma_s^2} + \alpha_y^* \big)$ to $\mathbb{I}_S \cap \mathbb{I}_x$ and $\mathbb{I}_S \cap \mathbb{I}_y$, respectively.

The gist of this corollary is that a mechanism can be designed to achieve the bound of Theorem 3.4 to each dimension independently and released with still-meaningful privacy guarantees. The reason is that this still includes all secret pairs $\mathcal{S}_{\text{pairs}}$

Proof. By independence, $X_{\mathbb{I}_x}$ and $X_{\mathbb{I}_y}$ can be treated as two unconnected traces of the type seen in **Figure 1**. As such the privacy guarantee of Theorem 3.4 can be upheld for each. The question is whether bounding CIP loss to the one-dimensional basic or compound secret associated with secret sets $\mathbb{I}_S \cap \mathbb{I}_x$ and $\mathbb{I}_S \cap \mathbb{I}_y$ still provides guarantees for the full secret set $\mathbb{I}_S$.

Without loss of generality, we will demonstrate for a basic and a compound secret. Consider the basic secret set $\mathbb{I}_S = \{X_{10}, X_{11}\}$, where $\mathbb{I}_S \cap \mathbb{I}_x = \{X_{10}\}$ (latitude) and $\mathbb{I}_S \cap \mathbb{I}_y = \{X_{11}\}$ (longitude). We again assume that independent gaussian noise of variance σ_s^2 is added to all $X_{\mathbb{I}_S}$, since this is optimal under utility constraints. We have now bounded the Rényi divergence when conditioning on pairs of hypotheses on latitude and longitude separately.

$$\mathcal{S}_{\text{pairs}_x} = \mathcal{S}_{\text{pairs}_y} = \{(x_s, x_s') : x_s \in \mathbb{R}, \|x_s - x_s'\|_2 \leq r\}$$

By independence, this also bounds the Rényi divergence conditioning on pairs of hypotheses on latitude and longitude jointly:

$$\mathcal{S}_{\text{pairs}_{xy}} = \{(x_s, x_s') : x_s \in \mathbb{R}^2, \|x_s - x_s'\|_2 \leq r\}$$

In effect, we have guaranteed privacy for any pair of hypotheses (s_i, s_j) in the square circumscribing the circle of radius r that we with to provide. The analysis on the direct privacy loss is exactly the same as it was in the more general case. Since the Rényi divergences of $X_{\mathbb{I}_U} \cap X_{\mathbb{I}_x}$ and of $X_{\mathbb{I}_U} \cap X_{\mathbb{I}_y}$ add, the α^*'s add.

The same goes for a compound secret. Consider three location compound secret pairs given by

$$\mathcal{S}_{\text{pairs}_{xy}} = \Big\{ \big(\{x_{s1}, x_{s2}, \dots\}, \{x_{s1}', x_{s2}', \dots\} \big) : x_{si} \in \mathbb{R}^2, \|x_{sk} - x_{sk}'\|_2 \leq r, \forall\ k \Big\}$$

Instead, we bound privacy loss for

$$\mathcal{S}_{\text{pairs}_x} = \mathcal{S}_{\text{pairs}_y} = \Big\{ \big(\{x_{s1}, x_{s2}, \dots\}, \{x_{s1}', x_{s2}', \dots\} \big) : x_{si} \in \mathbb{R}, \|x_{sk} - x_{sk}'\|_2 \leq r, \forall\ k \Big\}$$

Separately, giving us α_x^* and α_y^*. This again includes any two hypotheses on the three locations such that each pair of x_{sk}, x_{sk}' is within a square circumscribing a circle of radius r. We achieve this by bounding privacy loss for all Δ_x in a 3d L_2 ball of radius $\sqrt{S}\ r$, as with Δ_y.

This corollary can be extended to all traces of all dimensions that are probabilistically independent. □

We make use of the above proof in the Experiments section.

7.3 Derivation of Algorithms

In this section, we derive the three SDP-based algorithms of Section 4 and their properties.

7.3.1 Derivation of SDP_A

SDP_A minimizes the privacy loss bound of Theorem 3.4 for any compound or basic secret encoded by secret set $\mathbb{I}_S$. As is clarified in its proof (Appendix 7.2.4), the bound is tight when $\mathbb{I}_S$ encodes a basic secret. If $\mathbb{I}_S$ encodes a compound secret, the tightness depends on the conditional prior class Θ.

Our variable for minimizing this bound is the noise covariance matrix $\Sigma^{(g)}$. Due to the conditional independence exhibited by Lemma 3.2, $G_{\mathbb{I}_S}$ and $G_{\mathbb{I}_U}$ may be independent. The additive noise $G_i \in G_{\mathbb{I}_S}$ are all independent Gaussian with variance σ_s^2. This is because — conditioning on $\{X_{\mathbb{I}_S} = x_s\}$ — $Z_{\mathbb{I}_S}$ is independent of $X_{\mathbb{I}_U}$ and $Z_{\mathbb{I}_U}$. So, $G_{\mathbb{I}_S} \sim \mathcal{N}(\mathbf{0}, \sigma_s^2 I)$, and $\Sigma^{(g)}_{ss} = \sigma_s^2 I$. The additive noise $G_i \in G_{\mathbb{I}_U}$ are all dependent as described by $\Sigma^{(g)}_{uu}$, and $G_{\mathbb{I}_U} \sim \mathcal{N}(\mathbf{0}, \Sigma^{(g)}_{uu})$. Consequently, $\Sigma^{(g)}$ is completely characterized by $\Sigma^{(g)}_{uu}$ and σ_s^2.

To see how the bound of Theorem 3.4 can be redrafted as an SDP, first notice that its two terms may be written as the maximum eigenvalue of a matrix product. Here, $\Sigma_{\text{eff}} = A^\intercal B A$, where $A = \Sigma_{us}\Sigma_{ss}^{-1}$ and $B = \big(\Sigma_{u|s} + \Sigma^{(g)}_{uu}\big)^{-1}$

$$\frac{1}{\sigma_s^2} + \alpha^* = \text{maxeig}\big(\frac{1}{\sigma_s^2} I + A^\intercal B A\big) = \text{maxeig}\left(\begin{bmatrix} I & A \end{bmatrix} \begin{bmatrix} \frac{1}{\sigma_s^2} I & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} I \\ A \end{bmatrix} \right) = \text{maxeig}\big(\tilde{A}^\intercal \tilde{B} \tilde{A}\big)$$

This expression uses all parameters of $\Sigma^{(g)}$: σ_s^2 parametrizes $\Sigma^{(g)}_{ss}$ and $\Sigma^{(g)}_{uu} = B^{-1} - \Sigma_{u|s}$, where $\Sigma_{u|s}$ is given by the kernel function of Θ.

Before casting this as an SDP, we provide a formal definition from Vandenberghe & Boyd (1996):

Definition 7.1. *Semidefinite Program* The problem of minimizing a linear function of a variable $x \in \mathbb{R}^n$ subject to a matrix inequality:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \; & c^\intercal x \\ \text{s.t. } & F_0 + \sum_{i=1}^n x_i F_i \succeq 0 \\ & Ax = b \end{aligned}$$

where the $F_i \in \mathbb{R}^{n \times n}$ are all symmetric and $A \in \mathbb{R}^{p \times n}$ is a *semidefinite program*, or SDP.

The task of minimizing $\text{maxeig}\big(\tilde{A}^\intercal \tilde{B} \tilde{A}\big)$ under MSE constraints can almost be formulated as an SDP:

$$\begin{aligned} \min_{B \succeq 0, 1/\sigma_s^2 \geq 0} \; & \beta^* \\ \text{s.t. } & \beta^* I \succeq \tilde{A}^\intercal \tilde{B} \tilde{A} \\ & B \preceq \Sigma_{u|s}^{-1} \\ & \mathbf{tr}(\Sigma^{(g)}_{uu}) + |\mathbb{I}_S| \sigma_s^2 \leq n o_t \end{aligned}$$

Here, the first constraint guarantees that the maximum eigenvalue of $\tilde{A}^\intercal \tilde{B} \tilde{A}$ is bounded by β^*, which the objective minimizes. At program completion, we set $\Sigma^{(g)}_{uu} = B^{-1} - \Sigma_{u|s}$, and the second constraints ensures that this is still PSD. The final constraint bounds the MSE of the mechanism $\Sigma^{(g)}$. Note that $\mathbf{tr}(\Sigma^{(g)}_{uu}) + |\mathbb{I}_S|\sigma_s^2 = \mathbf{tr}(\Sigma^{(g)})$. The trouble lies the last constraint. Our program variable is B, but the final linear constraint requires $\Sigma^{(g)}$, which is expressed using the inverse of B. This is not immediately available in the SDP framework.

To make the final linear constraint available, we invert the above program using the observation that the maximum eigenvalue of $\tilde{A}^\intercal \tilde{B} \tilde{A}$ is the inverse of the minimum eigenvalue of $(\tilde{A}^\intercal \tilde{B} \tilde{A})^{-1}$. Instead of optimizing over B and $1/\sigma_s^2$, we optimize over B^{-1} and σ_s^2. Since $B^{-1} = \Sigma_{u|s} + \Sigma^{(g)}_{uu}$, we may now have a utility constraint directly on the trace of $\Sigma^{(g)}$. To make B^{-1} our program variable, we approximate $(\tilde{A}^\intercal \tilde{B} \tilde{A})^{-1}$ with $\tilde{A}^{-1} \tilde{B}^{-1} \tilde{A}^{-\intercal}$. First note that $\tilde{A} \in \mathbb{R}^{n \times |\mathbb{I}_S|}$, and has full column rank for the covariances we work with. So, $\tilde{A}^{-1} = (\tilde{A}^\intercal \tilde{A})^{-1} \tilde{A}^\intercal \in \mathbb{R}^{(|\mathbb{I}_S| \times n)}$ is the

left inverse of $\tilde{A}$ and is the least squares solution to $\tilde{A}^{-1}\tilde{A} = \tilde{A}^{\intercal}\tilde{A}^{-\intercal} = I$ (we denote its transpose as $\tilde{A}^{-\intercal}$). It is also the least squares solution to $\tilde{A}\tilde{A}^{-1} = \tilde{A}^{-\intercal}\tilde{A}^{\intercal} = I$. Thus, we have an approximation of the inverse $(\tilde{A}^{\intercal}\tilde{B}\tilde{A})^{-1}$:

$$\begin{aligned}(\tilde{A}^{\intercal}\tilde{B}\tilde{A})\,(\tilde{A}^{-1}\tilde{B}^{-1}\tilde{A}^{-\intercal}) &\approx \tilde{A}^{\intercal}\tilde{B}\tilde{B}^{-1}\tilde{A}^{-\intercal}\\ &= \tilde{A}^{\intercal}\tilde{A}^{-\intercal}\\ &\approx I\end{aligned}$$

We now can optimize in terms of B^{-1} with the augmented matrix $\tilde{B}^{-1}$:

$$\tilde{B}^{-1} = \begin{bmatrix} \sigma_s^2 I & 0 \\ 0 & B^{-1}\end{bmatrix}$$

We then optimize the following SDP:

$$\begin{aligned}\max_{B^{-1}\succeq 0,\sigma_s^2\geq 0}\ & \beta^*\\ \text{s.t. }& \beta^* I \preceq \tilde{A}^{-1}\tilde{B}^{-1}\tilde{A}^{-\intercal}\\ & B^{-1}\succeq \Sigma_{u|s}\\ & \mathbf{tr}(\tilde{B}) - \mathbf{tr}(\Sigma_{u|s}) \leq no_t\end{aligned}$$

Upon program completion we recover σ_s^2 and $\Sigma_{uu}^{(g)} = B^{-1} - \Sigma_{u|s}$ which we know is PSD due to the second constraint. The first constraint guarantees that the minimum eigenvalue of the approximated inverse is $\geq \beta^*$, which the objective maximizes. If the minimum eigenvalue of the approximate inverse is close to that of the true inverse, then we successfully minimize the maximum eigenvalue of $\tilde{A}^{\intercal}\tilde{B}\tilde{A}$, and thus minimize the direct and indirect privacy loss. The third constraint limits the MSE of $\Sigma^{(g)}$ since $\mathbf{tr}(\tilde{B}) - \mathbf{tr}(\Sigma_{u|s}) = (\mathbf{tr}(\Sigma_{uu}^{(g)}) + |\mathbb{I}_S|\sigma_s^2 + \mathbf{tr}(\Sigma_{u|s})) - \mathbf{tr}(\Sigma_{u|s}) = \mathbf{tr}(\Sigma^{(g)})$. By inverting $\tilde{A}^{\intercal}\tilde{B}\tilde{A}$, this constraint is available in the SDP framework.

By expressing the above program in terms of the variable $\Sigma^{(g)}$ instead of indirectly via B^{-1} and σ_s^2, we get SDP_A:

$$\begin{aligned}\mathbf{SDP_A}:\quad \arg\max_{\Sigma^{(g)}\succeq 0}\ & \beta^*\\ \text{s.t. }& \tilde{A}^{-1}\tilde{B}^{-1}\tilde{A}^{-\intercal} \succeq \beta^*\mathbf{I}\\ & \mathbf{tr}(\Sigma^{(g)}) \leq no_t\end{aligned}$$

It is straightforward to write this SDP in the form seem in Definition 7.1. The program variables x would be the diagonal and upper or lower triangular part of $\Sigma^{(g)}$ along with β^*. With some linear algebra, the first constraint can be written in the form of $F_0 + \sum_{i=1}^n x_i F_i \succeq 0$, and the second constraint can be written as $Ax = b$. With the use of contemporary convex programming tools like CVXOPT (Vandenberghe, 2010) rewriting into this form is unnecessary.

7.3.2 Derivation of SDP_B

SDP_B takes a set of covariance matrices $\mathcal{F} = \{\Sigma_1, \ldots, \Sigma_k\}$, each of which is designed to protect some secret set $\mathbb{I}_{Si}$, and returns a covariance matrix $\Sigma^{(g)}$ that preserves the privacy loss bound of each Σ_i to each $\mathbb{I}_{Si}$. It does so while minimizing the utility loss of $\Sigma^{(g)}$. This algorithm is also expressed as an SDP. It is based on the following corollary, which we have omitted from the main text:

Corollary 7.2.2. More PSD, More Private: *For a basic or compound secret denoted by indices $\mathbb{I}_S$, the CIP loss bound of Equation 5 provided by a Gaussian noise mechanism with covariance $\Sigma^{(g)}$ is lower than it would be for any $\Sigma^{(g)\prime} \prec \Sigma^{(g)}$.*

Proof. First note that if $\Sigma^{(g)} \succ \Sigma^{(g)\prime}$, then the same is true for its sub-matrices:

$$\Sigma_{ss}^{(g)} \succ \Sigma_{ss}^{(g)\prime} \qquad \Sigma_{uu}^{(g)} \succ \Sigma_{uu}^{(g)\prime}$$

Recall the privacy loss bound of Equation 5:

$$\varepsilon \leq \frac{\lambda}{2} S r^2 \Big(\frac{1}{\sigma_s^2} + \alpha^* \Big)$$

Also recall that $\Sigma_{ss}^{(g)} = \sigma_s^2 I$ and $\Sigma_{ss}^{(g)'} = \sigma_s^{2\prime} I$. Since $\Sigma_{ss}^{(g)} \succ \Sigma_{ss}^{(g)'}$, we already know that $\sigma_s^2 > \sigma_s^{2\prime}$, and thus the first term of Equation 5 is lower for $\Sigma^{(g)}$.

It remains to show that the second term is also lower, $\alpha^* < \alpha^{*\prime}$. Starting with what we're given,

$$\begin{gathered}
\Sigma_{uu}^{(g)} \succ \Sigma_{uu}^{(g)'} \\
\Sigma_{uu}^{(g)} + \Sigma_{u|s} \succ \Sigma_{uu}^{(g)'} + \Sigma_{u|s} \\
(\Sigma_{uu}^{(g)} + \Sigma_{u|s})^{-1} \prec (\Sigma_{uu}^{(g)'} + \Sigma_{u|s})^{-1} \\
B \prec B' \\
A^\intercal B A \prec A^\intercal B' A \\
\max \operatorname{eig}(A^\intercal B A) < \max \operatorname{eig}(A^\intercal B' A) \\
\alpha^* < \alpha^{*\prime}
\end{gathered}$$

Therefore $\frac{1}{\sigma_s^2} + \alpha^* < \frac{1}{\sigma_s^{2\prime}} + \alpha^{*\prime}$, and the CIP bound of Equation 5 is lower for $\Sigma^{(g)}$ than it is for $\Sigma^{(g)'}$. □

With Corollary 7.2.2 in mind, SDP_B is natural:

$$\begin{aligned}
\textbf{SDP}_\textbf{B}: \quad & \arg\min_{\Sigma^{(g)}} \mathbf{tr}(\Sigma^{(g)}) \\
& \text{s.t. } \Sigma^{(g)} \succeq \Sigma_i^{(g)},\ \forall \Sigma_i^{(g)} \in \mathcal{F}
\end{aligned}$$

SDP_B attempts to minimize, but does not constrain, the utility loss of the chosen $\Sigma^{(g)}$. To provide an upper bound on the resulting utility loss, we provided the following claim in the main text:

Claim Utility loss of SDP_B: *The utility loss of $\Sigma^{(g)} = SDP_B(\mathcal{F})$ is no greater than $\sum_{\Sigma_i \in \mathcal{F}} \mathbf{tr}(\Sigma_i)$.*

Proof. The covariance $\Sigma^{(g)'} = \sum_{\Sigma_i^{(g)} \in \mathcal{F}} \Sigma_i^{(g)}$ with MSE $\sum_{\Sigma_i^{(g)} \in \mathcal{F}} \mathbf{tr}(\Sigma_i^{(g)})$ is in the feasible set of SDP_B problem since $\Sigma^{(g)'} \succeq \Sigma_i^{(g)},\ \forall \Sigma_i^{(g)} \in \mathcal{F}$. Unless $\Sigma^{(g)'}$ has the lowest MSE of all $\Sigma^{(g)}$ in the feasible set, a covariance matrix with better utility will be chosen. □

7.3.3 Derivation of Algorithm 1, Multiple Secrets

Multiple Secrets combines SDP_A and SDP_B to minimize the privacy loss to each basic secret within a trace. The basic mechanism is useful in cases when inferences at each time within the trace — each basic secret — is sensitive.

Let $\mathbb{I}_{Si}$ be the secret set representing basic secret i, of which there are N (e.g. if location is sampled at N times). Then $\mathbb{I}_{\mathcal{S}_b} = \{\mathbb{I}_{S1}, \ldots, \mathbb{I}_{SN}\}$ contains the indices corresponding to each. Multiple Secrets works by first producing N covariance matrices, $\Sigma_i^{(g)} = \text{SDP}_\text{A}(\mathbb{I}_{Si}, \Sigma, o_t)$ on each basic secret. It then uses $\text{SDP}_\text{B}(\mathcal{F} = \{\Sigma_1^{(g)}, \ldots, \Sigma_N^{(g)}\})$ to produce a single covariance matrix $\Sigma^{(g)}$ that preserves the privacy loss to each basic secret (note that, being basic secrets, the privacy loss bound that SIG OPT optimizes is tight).

By virtue of using SDP_B, the MSE of the resultant $\Sigma^{(g)}$ is minimized but not constrained. To bound the MSE of the Basic Mechanism by O, we may simply bound the MSE of each $\Sigma_i^{(g)}$ by $o_t = {}^{O}/_{N}$. Then, by the above Claim, the MSE of the solution cannot be greater than O. In practice, this bound may be too loose. We hope to tighten it in future work.

7.4 Experimental details

We use a 2d location trace and a 1d home temperature dataset. For the location data, having observed that the correlation between latitude and longitude is low (≈ 0.06) we treat each dimension as independent. By way of Corollary 7.2.1, this allows us to bound privacy loss and design mechanisms for each dimension separately. Furthermore, having observed that each dimension fits the nearly the same conditional prior, we treat our dataset of 10k 2-dimensional traces as a dataset of 20k 1-dimensional traces, where each trace represents one dimension of a 2d location trajectory.

The one-dimensional traces of temperature and location are indexed by timestamps, for which we would use the following kernel functions:

$$k_{\text{RBF}}(t_i, t_j) = \sigma_x^2 \exp\Big(-\frac{(t_i - t_j)^2}{2l^2}\Big) \qquad k_{\text{PER}}(t_i, t_j) = \sigma_x^2 \exp\Big(\frac{-2\sin^2(\pi|t_i - t_j|/p)}{l^2}\Big) \tag{6}$$

to determine the covariance between two points sampled at times t_i and t_j. The parameters including variance σ_x^2 and length scale l. The lengthscale determines the window of time in which two sampled points are highly correlated.

Preprocessing of location data We first limit the dataset to traces of under 50 locations that are between 4.5 and 5.5 minutes in duration. Caring only about the conditional dependence between locations, we then de-mean each trace and normalize its variance to one. Normalizing the variance of traces implicitly sets $\sigma_x^2 = 1$ in the above RBF kernel, in essence assuming that the adversary has a decent prior for the user's average speed in a given trace, and could do the same operation.

Fitting of location data We then find the maximum likelihood RBF kernel for each distinct trace. Having fixed the variance σ_x^2, this amounts to fitting only the length scale for each dimension, l_x and l_y, individually. The length scale represents the average window of time during which neighboring locations are highly correlated (i.e. correlation > 0.8). Relatively smooth traces will have large length scales and chaotic traces will have low length scales. However, the fact that sampling rates vary significantly between traces means that traces with equal length scales can have very different degrees of correlation. To encapsulate both of these effects, we study the empirical distribution of *effective* length scale of each trace

$$l_{\text{eff},x} = \frac{l_x}{P} \quad l_{\text{eff},y} = \frac{l_y}{P}$$

where P is the trace's sampling period and l_x, l_y are the its optimal length scales. $l_{\text{eff},x}, l_{\text{eff},y}$ tell us the average number of neighboring locations that are highly correlated, instead of time period. For instance, a given trace with an optimal $l_{\text{eff},x} = 8$ tells us that every eight neighboring location samples in the x dimension have correlation > 0.8. The empirical distribution of effective length scales across all traces describes – over a range of logging devices (sampling rates), users, and movement patterns – how many neighboring points are highly correlated in location trace data. After this preprocessing, we are able to use the kernels that take indices (not time) as arguments.

$$k_{\text{RBF}}(i, j) = \exp\Big(-\frac{(i - j)^2}{2l_{\text{eff}}^2}\Big) \qquad k_{\text{PER}}(i, j) = \exp\Big(\frac{-2\sin^2(\pi|i - j|/p)}{l_{\text{eff}}^2}\Big)$$

In each plot we then observed a spectrum of conditional priors by sweeping the effective length scale and plotting posterior uncertainty for various noise mechanisms of equal utility loss. This ranges from a prior assuming nearly independent location samples (chaotic trace) on the left up to highly dependent location samples (traveling in a straight line or standing still) on the right. To understand how realistic these conditional prior parameters are, we displayed the middle 50% of the empirical distribution of l_{eff} (x and y together) from the GeoLife dataset. Note that the distribution of l_{eff_x} and l_{eff_y} are nearly identical.

To compute posterior uncertainty, we consider a 50-point one-dimensional location trace. The basic secret is a single index in the middle of the trace, and the compound secret consists of two neighboring indices also in

the middle of trace. For each value of l_{eff}, we compute the $\mathbb{R}^{50\times 50}$ conditional prior covariance matrix Σ using the RBF kernel above. We then compare the posterior uncertainty when $\Sigma^{(g)}$ is an Approach C baseline, or an optimized covariance matrix using one of the three algorithms. We re-optimize $\Sigma^{(g)}$ for each l_{eff}, since each l_{eff} represents a different conditional prior class. The MSE is fixed in all figures except the two exhibiting "All Basic Secrets", where SDP_B is used. Recall that this algorithm minimizes utility loss while maintaining a series of privacy guarantees. Here, the MSE is identical across mechanisms for each l_{eff}, but changes from one l_{eff} to another.

For the temperature data, our preprocessing steps were nearly identical, except we use the periodic kernel instead of the RBF kernel, and we did not need to remove any traces from the dataset, as the data was much cleaner.

Computation of Posterior Uncertainty Interval Each of the plots in **Figure 2** shows the 2σ uncertainty interval on $X_{\mathbb{I}_S}$ of a Gaussian process Bayesian adversary with prior covariance Σ and any mean function

The posterior covariance is computed using standard formulas for linear Gaussian systems. Knowing that $Z = X + G$, we may write the joint precision matrix Λ (inverse of covariance matrix) of (X, Z) as

$$\Lambda^{(X,Z)} = \begin{bmatrix} \Sigma^{-1} + {\Sigma^{(g)}}^{-1} & -{\Sigma^{(g)}}^{-1} \\ -{\Sigma^{(g)}}^{-1} & {\Sigma^{(g)}}^{-1} \end{bmatrix}$$

It is then a well known result that the conditional covariance matrix is given by

$$\begin{aligned} \Sigma_{x|z} &= \Lambda_{xx}^{-1} \\ &= \left(\Sigma^{-1} + {\Sigma^{(g)}}^{-1}\right)^{-1} \end{aligned}$$

This provides the posterior covariance of all locations X given any released trace Z that uses a Gaussian mechanism with covariance $\Sigma^{(g)}$. Note that the CIP guarantee naturally keeps posterior uncertainty large since the posterior density at any two x_s close together must be similar. For these Gaussian posteriors, 2σ tells us the adversary's 68% confidence interval on $X_{\mathbb{I}_S}$ after obvserving Z.

For basic secrets (one location), we simply report twice the posterior standard deviation at the sensitive index i, given by

$$2\sqrt{\Sigma_{x|z,ii}}\ .$$

For compound secrets involving multiple locations the posterior distribution is a length $|\mathbb{I}_S|$ multivariate normal with covariance $\Sigma_{x|z,ss}$. Intuitively, we wish to find the direction of the vector $X_{\mathbb{I}_S}$ in which the posterior interval is the *shortest*. This is the worst case posterior interval on the compound secret. We do this by reporting

$$2\sqrt{\text{mineig }\Sigma_{x|z,ss}}\ .$$

7.5 Discussion of GP Conditional Prior Class

Recall that a conditional prior class requires for any $P_{\mathcal{P}_i}, P_{\mathcal{P}_j} \in \Theta$ that

$$P_{\mathcal{P}_i}(X_{\mathbb{I}_U}|X_{\mathbb{I}_S} = x_s) = P_{\mathcal{P}_j}(X_{\mathbb{I}_U} + c^u_{ij\mathbb{I}_S}|X_{\mathbb{I}_S} = x_s + c^s_{ij\mathbb{I}_S})$$

for all x_s. Notice that the mapping $(x_s, x'_s) + c^s_{ij\mathbb{I}_S}$ is a bijection from $\mathcal{S}_{\text{pairs}}$ onto itself. As such, each pair of conditional distributions,

$$\left(P_{\mathcal{P}_j}(X_{\mathbb{I}_U}|X_{\mathbb{I}_S} = x_s), P_{\mathcal{P}_j}(X_{\mathbb{I}_U}|X_{\mathbb{I}_S} = x'_s)\right)$$

induced by $(x_s, x'_s) \in \mathcal{S}_{\text{pairs}}$ is a mean-shifted version of the pair of distributions

$$\left(P_{\mathcal{P}_i}(X_{\mathbb{I}_U}|X_{\mathbb{I}_S} = x_s - c^s_{ij\mathbb{I}_S}), P_{\mathcal{P}_i}(X_{\mathbb{I}_U}|X_{\mathbb{I}_S} = x'_s - c^s_{ij\mathbb{I}_S})\right)$$

induced by $(x_s, x'_s) - c^s_{ij\mathbb{I}_S} \in \mathcal{S}_{\text{pairs}}$. Since the Rényi divergence between two distributions and two mean-shifted versions thereof is unchanged, we may use one additive noise mechanism for all priors in class Θ.

To see how this applies to the GP prior class, recall the formula for a conditional multivariate Gaussian distribution:

$$P(X_{\mathbb{I}_U} | X_{\mathbb{I}_S} = x_s) = \mathcal{N}(\mu_{u|s}, \Sigma_{u|s})$$

where,

$$\begin{aligned}\mu_{u|s} &= \mu_u + \Sigma_{us}\Sigma_{ss}^{-1}(x_s - \mu_s) \\ \Sigma_{u|s} &= \Sigma_{uu} - \Sigma_{us}\Sigma_{ss}^{-1}\Sigma_{su}\end{aligned}$$

A GP prior class includes all GP distributions with a fixed kernel $k(t_i, t_j)$ and any mean function $\mu(t)$. For a fixed set of time points, this corresponds to a fixed covariance matrix Σ and any mean parameters $\boldsymbol{\mu}$:

$$X \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$$

Let $P_{\mathcal{P}_i} = \mathcal{N}(\bar{\boldsymbol{\mu}}, \Sigma)$ and $P_{\mathcal{P}_j} = \mathcal{N}(\hat{\boldsymbol{\mu}}, \Sigma)$, then conditioned on some sensitive points $X_{\mathbb{I}_S}$ the distribution on $X_{\mathbb{I}_U}$ has the same covariance $\Sigma_{u|s}$ and conditional means

$$\begin{aligned}\bar{\mu}_{u|s} &= \bar{\mu}_u + \Sigma_{us}\Sigma_{ss}^{-1}(x_s - \bar{\mu}_s) \\ &= (\bar{\mu}_u - \Sigma_{us}\Sigma_{ss}^{-1}\bar{\mu}_s) + \Sigma_{us}\Sigma_{ss}^{-1}x_s \\ \hat{\mu}_{u|s} &= \hat{\mu}_u + \Sigma_{us}\Sigma_{ss}^{-1}(x_s - \hat{\mu}_s) \\ &= (\hat{\mu}_u - \Sigma_{us}\Sigma_{ss}^{-1}\hat{\mu}_s) + \Sigma_{us}\Sigma_{ss}^{-1}x_s\end{aligned}$$

which implies that the conditional distributions are identical up to a mean shift for the *same* x_s value.

$$P_{\mathcal{P}_i}(X_{\mathbb{I}_U} | X_{\mathbb{I}_S} = x_s) = P_{\mathcal{P}_j}(X_{\mathbb{I}_U} + c^u_{ij\mathbb{I}_S} | X_{\mathbb{I}_S} = x_s)$$

for all x_s. Here, $c^u_{ij\mathbb{I}_S} = (\bar{\mu}_u - \Sigma_{us}\Sigma_{ss}^{-1}\bar{\mu}_s) - (\hat{\mu}_u - \Sigma_{us}\Sigma_{ss}^{-1}\hat{\mu}_s)$, and $c^s_{ij\mathbb{I}_S} = 0$.

To see how this allows a single additive mechanism to work for all mean functions, notice that we also have

$$P_{\mathcal{P}_i}(X_{\mathbb{I}_U} | X_{\mathbb{I}_S} = x'_s) = P_{\mathcal{P}_j}(X_{\mathbb{I}_U} + c^u_{ij\mathbb{I}_S} | X_{\mathbb{I}_S} = x'_s)$$

for x'_s, so the divergences

$$\begin{aligned}D_\lambda\begin{pmatrix}P_{\mathcal{P}_i}(X_{\mathbb{I}_U} | X_{\mathbb{I}_S} = x_s) \\ P_{\mathcal{P}_i}(X_{\mathbb{I}_U} | X_{\mathbb{I}_S} = x'_s)\end{pmatrix} &= D_\lambda\begin{pmatrix}P_{\mathcal{P}_j}(X_{\mathbb{I}_U} + c^u_{ij\mathbb{I}_S} | X_{\mathbb{I}_S} = x_s) \\ P_{\mathcal{P}_j}(X_{\mathbb{I}_U} + c^u_{ij\mathbb{I}_S} | X_{\mathbb{I}_S} = x'_s)\end{pmatrix} \\ &= D_\lambda\begin{pmatrix}P_{\mathcal{P}_j}(X_{\mathbb{I}_U} | X_{\mathbb{I}_S} = x_s) \\ P_{\mathcal{P}_j}(X_{\mathbb{I}_U} | X_{\mathbb{I}_S} = x'_s)\end{pmatrix}\end{aligned}$$

are equal. The same goes for the noisy trace $X_{\mathbb{I}_U} + Z_{\mathbb{I}_U} | X_{\mathbb{I}_S} = x_s$, when Z is drawn independently of X, allowing us to bound privacy loss for all $P \in \Theta$.