Dynamically Protecting Privacy, under Uncertainty

Mine Su Erturk
Graduate School of Business
Stanford University
mserturk@stanford.edu

Kuang Xu
Graduate School of Business
Stanford University
kuangxu@stanford.edu

Abstract

We propose and analyze the $\epsilon$-Noisy Goal Prediction Game to study a fundamental privacy versus efficiency tradeoff in dynamic decision-making under uncertainty. An agent wants to quickly reach a final goal in a network through a sequence of actions, while the effects of these actions are subject to random noise and perturbation. Meanwhile, an overseeing adversary observes the effects of the agent’s past actions and tries to predict the goal. We are interested in understanding the probability that the adversary predicts the goal correctly (prediction risk) as a function of the time it takes the agent to reach her goal (delay).

Our main results characterize the prediction risk versus delay tradeoff under various network topologies. First, we establish an asymptotically tight characterization in complete graphs, showing that (1) intrinsic uncertainty always leads to a strictly positive overhead for the agent, even as her delay tends to infinity, and (2) under a carefully designed decision policy the overhead can be only additive with respect to the noise level, and thus incur an asymptotically negligible effect on system performance. We further apply these insights to studying network topologies generated by random graphs, and to designing private networks. In both cases, we show how to achieve an additive overhead even for relatively sparse, non-complete networks. Finally, for general graphs, we construct a private agent strategy that can operate under any level of intrinsic uncertainty. Our analysis is centered around a new class of “noise-harvesting” agent strategies which adaptively combines intrinsic uncertainty with additional artificial randomization to achieve efficient obfuscation.\footnote{Draft version: October 2019.}

Keywords: goal-oriented privacy, dynamic decision-making, stochastic shortest path, algorithms, probability.

1 Introduction

The advancement in machine learning and data collection infrastructures has made it increasingly easy for firms, e-commerce platforms, or governmental entities to collect and analyze the behavior and actions of individuals or competitors (Valentino-DeVries et al. (2018), Mayer et al. (2016), De Montjoye et al. (2013)). Such analysis enables the entity to make powerful predictions on sensitive information that an individual under monitoring would not like to reveal, such as future behaviors or preferences. Motivated by the threats to privacy posed by these developments, a number of papers in recent years have been devoted to designing privacy-preserving policies for dynamic decision-making problems (cf. Fanti...
et al. (2015), Tossou and Dimitrakakis (2016)). As a general theme, the decision maker would introduce carefully curated, additional randomization in her actions (compared to a privacy-oblivious optimal policy), so that the randomness will render it provably difficult for a third-party to accurately predict the sensitive information based on the observed actions alone.

We aim to contribute to the understanding of privacy-preserving dynamic decision-making, by focusing on problems with intrinsic uncertainty, a key feature that has thus far garnered little attention. The majority of existing work assumes perfect control, where the decision maker’s action maps to changes in the system in a deterministic manner. Examples of such approaches include the state updates in Tsitsiklis and Xu (2018), query locations in Xu (2018), Tsitsiklis et al. (2018), epidemic infections in Fanti et al. (2015), and manipulations to entries of a database in Dwork et al. (2014). In contrast, we will focus on problems with a non-trivial amount of uncertainty, so that the effects of any action are subject to stochastic perturbation and noise, over which the decision maker has no control.

There are two main reasons for going beyond models with perfect control to incorporate uncertainty. On a practical level, noise and uncertainty are inherent in a vast array of real-world systems of dynamic decision-making, such as routing (Flajolet et al. (2017), Jaillet et al. (2016)), navigation (Sezer et al. (2015)) and planning (Mestre et al. (2015)); models that assume perfect control fail to capture these applications. On a conceptual level, crafting efficient decision policies under uncertainty requires a fundamentally different set of tools and design philosophies that differ substantially from their perfect-control counterparts. For instance, as we will show in this paper, seemingly natural extensions of policies designed under the perfect-control assumptions can perform significantly more poorly under uncertainty.

ε-Noisy Goal Prediction Game. The present paper takes a first step towards building a framework that will allow us to design and analyze privacy-preserving policies in dynamic decision-making problems under uncertainty. We propose an ε-Noisy Goal Prediction Game, and study the optimal tradeoff between delay and privacy facing the decision maker under different network topologies. The ε-Noisy Goal Prediction Game serves as a privacy-aware version of the classical, privacy-oblivious stochastic shortest path problem (cf. Bertsekas and Tsitsiklis (1991), Eaton and Zadeh (1962), Croucher (1978)). As such, it enjoys a wide range of applications in routing (cf. Flajolet et al. (2017), Kim et al. (2005)), evacuation planning (Wang et al. (2016)), and transportation (cf. Miller-Hooks (2001)). We will discuss some of the motivating applications in Section 1.1.

We begin with an informal description of the model. The decision maker is an agent who operates in discrete time and whose state at time $t \in \mathbb{Z}^+$ corresponds to a vertex in an undirected graph $G = (\mathcal{V}, \mathcal{E})$. Starting from some initial vertex, the agent’s main objective is to reach a goal vertex, $D$, as quickly as possible; by traversing along the edges of the graph $G$. The goal vertex is generated uniformly at random from $\mathcal{V}$ at $t = 0$, and its identity is unknown to the adversary. Crucially, the agent does not have perfect control over her trajectory, and her movements over the graph are subject to uncertainty as follows: in each time period, the agent may choose a vertex, $v$, among those connected to her current state as her action, and in the following period, her state will become $v$ with probability $1 - \varepsilon$, and be set to a random neighboring vertex, otherwise. That is, the effect of the agent’s action coincides with her true intention with probability $1 - \varepsilon$, and is otherwise random.

Meanwhile, an overseeing adversary monitors all of the effects of the agent’s actions (i.e., he observes the agent’s trajectory across the graph), whose objective is to predict the goal $D$ before it is reached by the agent. The adversary has one opportunity at guessing the goal
but can exercise this opportunity at a time of his choosing. The adversary wins the game if he correctly guesses $D$ by the time it is reached by the agent, and loses, otherwise.

**Delay vs. Prediction Risk Tradeoff.** Being aware of the negative repercussions should her goal be predicted, the agent seeks a decision strategy that will minimize the probability of the goal being predicted by the adversary, $q$, subject to an upper bound on the average time for reaching the goal, $w$. We refer to $w$ as the agent’s delay, and $q$ the resulting **prediction risk.** In general, delay captures the agent’s cost, and prediction risk her level of privacy. The agent thus faces a fundamental tradeoff: on the one hand, the agent needs to incorporate intentional randomization into her actions, as greedily minimizing her delay could result in early trajectories being too revealing of her final goal and thus elevate the prediction risk; on the other hand, any randomization risks taking her further away from the goal, thus prolong the delay. Our main goal is to understand this tradeoff between delay and prediction risk, and furthermore, to identify efficient strategies that provably achieve, or perform close to, the optimal tradeoff.

Our model is inspired by and builds on the Goal Prediction Game proposed by Tsitsiklis and Xu (2018), by incorporating the new dimension of uncertainty. The original model was conceived as a privacy-aware variant of the deterministic shortest path problem, and establishes upper and lower bounds on the delay-predictability tradeoff; it corresponds to the special case of $\varepsilon = 0$ (noiseless) in our formulation.

### 1.1 Motivating Examples

While stylized, the $\varepsilon$-Noisy Goal Prediction Game captures the strategic considerations in applications where a privacy-conscious decision maker strives to reach a goal in a discrete manner, while coping with uncertainty in the effects of her actions. We illustrate in this section some of these motivating applications in operations management, transportation and communication networks, where uncertainty plays a notable role. The reader is also referred to the applications discussed in Tsitsiklis and Xu (2018) for the noiseless Goal Prediction Game, which falls within our framework as a special case.

#### 1.1.1 Anonymous Communication and Web-Browsing

Anonymous messaging protocols such as Bitmessage (Warren (2012)), Riposte (Corrigan-Gibbs et al. (2015)), and Herbivore (Goel et al. (2003)) enable users to send messages without revealing the identity of the receiver. A number of platforms, including the three mentioned above, ensure anonymity by broadcasting each message to a large number of nodes in the network, while only the recipient node has the private key to decode the message. However, uniformly broadcasting the message to the entire network can be costly as the network grows, or when the underlying physical communication network is sparse, e.g., when some nodes are geographically separated from each other. Hence, there has been growing interest in the community to design more efficient policies to ensure anonymous communication.

The Goal Prediction Game can be viewed as a more general model of anonymous messaging where the message is passed along the edges of the underlying communication network: the sender may pass the message to one of her neighbors, who subsequently relays it to one of his neighbors and so on. More concretely, the agent in the anonymous messaging scenario corresponds to the message that is being routed over a communication network. The vertices model the users on the communication network whereas the edges denote the possible connections between users that can be utilized by the routing protocol. Thus, we
can describe the path followed by the message using the trajectory and model the receiver of the message as the goal vertex $D$. Further, the adversary represents anyone with the ability to monitor the traffic over the network such as advertisers, state and government agencies.

Moreover, there can be various sources of uncertainty in these systems. For instance, certain nodes may forward the message to a random neighbor with positive probability instead of the actual vertex prescribed by the protocol, due to malicious attacks (e.g., routing attacks in Vu et al. (2009)) or malfunctioning. Alternatively, the links in the underlying communication network may change dynamically over time and a message may not be able to be passed onto an intended neighbor when that neighbor is offline or due to random link failures. As such, our model and algorithms can be used to aid the design of robust and efficient anonymous message routing protocols in these scenarios, and the delay in our model serves as a measure of the network traffic overhead. In Appendix D, we describe in detail the Bitmessage protocol and explain how our algorithms can potentially be implemented.

Secure web-browsing (cf. Dingledine et al. (2004)) can be modeled in a similar manner. In this scenario, instead of sending a message over the peer-to-peer network, the agent tries to access a particular website via proxy servers or VPN servers. While routing traffic through the network, the protocol sends the agent to a random server if the originally intended server is blocked with some probability. Once again, the agent wants to access the website as quickly as possible without revealing her intentions while the adversary is trying to identify the agent’s target website.

1.1.2 Secure Logistics in Supply Chain Operations

With the advent of machine learning and computer vision techniques, it has become increasingly easier for firms to collect and process data on their competitors’ operations. Hence, secure logistics in supply chain operations, where a firm wishes to ship a parcel from location A to B while keeping the identity of B as hidden from competitors, has become an emerging application. For instance, there has been growing evidence on firms using satellite imagery to track cars in parking lots (cf. The Economist (2019), Rodriguez (2018)), or ships approaching ports for delivery (Allioux (2018)). Given the wide availability of such techniques, it is plausible that a firm might track the frequency with which its competitor receives deliveries by observing the truck on its delivery route or by identifying when a loading dock is being used. Then, this estimated frequency could be used to make inferences about the sales volume or inventory of the competitor. Therefore, it is conceivable that a firm might find it useful to add obfuscation stops to its delivery route where the carrier may actually provide other services while pretending to be making deliveries.

Specifically, the agent in the secure logistics application represents the truck or the parcel that is being routed on a road network and the adversary is the competitor. The edges and the vertices in the graph describe the road infrastructure and the dropoff locations, respectively. Finally, unexpected disruptions such as weather conditions and political unrest often mandate the rerouting of the goods (Nagurney et al. (2018)), which can be modeled as noise or shocks that alter the agent’s intended movements.

1.1.3 Autonomous Vehicles

Autonomous vehicles are on track to becoming a primary mode of urban transportation in the future. However, there has also been increasing concern about potential privacy implications as a result of such automation. For instance, the recent news article by Ratnam (2019) reports how car manufacturers monitor the past trajectory of vehicles, which could
then be used to accurately predict its future destinations. It is therefore conceivable that, instead of following the exact shortest path between the origin and the destination, one may want the vehicles to incorporate small randomization into their route planning so as to obfuscate the destination at a minor cost of increased delay. Furthermore, if the traffic at a certain street or intersection becomes congested, then the vehicle may need to choose an alternative path which it could have not foreseen; such stochastic route variability can be captured by our formulation.

1.2 Summary of Main Contributions

We now give an informal preview of our main results, which characterize the prediction risk vs. delay tradeoff under different topologies with varying degrees of generality. The formal statements will be given in Section 3.

1. Tight Tradeoff in Complete Graphs (Theorem 1) We first consider the case where the graph $G$ is a complete graph, which, despite being a seemingly simple topology, turns out to contain a surprising amount of strategic richness that demands careful policy design. Furthermore, policies and analysis developed for complete graphs will serve as a foundation for subsequent analysis of non-complete graphs.

We provide a characterization of the tradeoff between delay and prediction risk, which is shown to be asymptotically tight as the graph size tends to infinity. Specifically, denote by $Q_p(w)$ the minimal prediction risk that can be achieved across all agent strategies with a delay of at most $w$. Fix noise level $\varepsilon < 1$, and delay target $w = \frac{\sqrt{\varepsilon}}{1+\varepsilon} + \frac{1}{2}$. We show that, as the graph size $n \to \infty$,

$$Q_p(w) = \frac{1}{2w-1} - \frac{1}{2w-1} \left(\frac{\varepsilon}{w - \beta_p(w)} + o(1)\right), \tag{1}$$

where $\beta_p(w)$ is a discrepancy term with $|\beta_p(w)| = O((\varepsilon/w)^2)$, and the term $o(1)$ tends to 0 as $n \to \infty$. As a part of the proof, we also provide a specific family of agent strategies, dubbed the Water-Filling Strategies, which asymptotically delivers the prediction risk given in Eq. (1).

Some important observations can be made from Eq. (1). First, for any noise level, the minimal prediction risk always scales inversely proportionally as a function of the delay. Second, the presence of uncertainty results in a strictly positive, additive efficiency overhead, whose magnitude is given by

$$\text{cost of uncertainty} = \frac{1}{2w-1} \cdot \frac{\varepsilon}{1 - \varepsilon} + O((\varepsilon/w)^2). \tag{2}$$

On the negative side, Eq. (1) shows that the presence of uncertainty always adversely affects the agent compared to a model with full control, even if the delay significantly surpasses the (stochastic) diameter of the graph (which is on the order of $O((1 - \varepsilon)^{-1})$). On the positive side, we show that the optimal tradeoff can be achieved by a carefully designed agent strategy which "harvests" the uncertainty intrinsic to the system and combines it with the artificial randomization that the strategy itself injects. As a result, the overhead on delay is only additive, and becomes increasingly negligible as $w$ grows. In summary, Eq. (1) shows that, at a qualitative level,

$$\text{prediction risk} = \Theta \left( \frac{1}{\text{delay} - \text{cost of uncertainty}} \right).$$
2. Designing Privacy-Friendly Network Topologies (Theorem 2) In the next two results, we apply the techniques developed for complete graphs to the study of non-complete topologies. First, we consider scenarios in which the agent has the option of designing the network structure beforehand (such as communication networks), and a complete graph topology requires such a high degree of connectivity that can be prohibitively expensive. It is therefore natural to ask: subject to a constraint on the average degree, what network topologies will best facilitate the deployment of privacy-aware agent strategies? Specifically, can a sparse network continue to deliver, just like in the case of a complete graph, an additive delay overhead in the presence of uncertainty?

We provide a positive answer to the above question, by adapting the agent strategy and analysis developed for complete graphs. For any degree sequence \( \bar{p}_n \) with \( \sqrt{n} \ll \bar{p}_n \leq n \), we show that there exists a family of graphs, \( G(n, \bar{p}_n) \), where each \( G \in G(n, \bar{p}_n) \) has \( n \) vertices and average degree at most \( \bar{p}_n \), over which the minimal prediction risk satisfies

\[
\frac{1}{2w+1} \leq Q(w) \leq \frac{1}{2w - 1 - \frac{\varepsilon}{(1-\varepsilon)^2}} + o(1),
\]

as \( n \to \infty \), whenever \( w > \frac{1}{2} + \frac{\sqrt{\varepsilon}}{1-\varepsilon} \). Note that the delay overhead due to uncertainty remains additive.

3. Non-Complete Graphs (Theorem 3) In various settings, however, the network is given exogenously and the agent does not have the option of changing its topology. Our next result examines a family of non-complete graphs that can be generated with high probability using the Erdős-Rényi random graph model, \( G(n, p) \). We design the Random Walk Water-Filling Strategy, which is applicable on any undirected graph, and use it to show that a graph from this family admits the minimal prediction risk

\[
\frac{1}{2w+1} \leq Q(w) \leq \frac{1}{2w - 2 - \frac{\varepsilon}{p(1-\varepsilon)^2}} + o(1),
\]

as \( n \to \infty \), whenever \( w > \left( \frac{1}{2} + \frac{\sqrt{\varepsilon}}{1-\varepsilon} \right) \frac{1}{p} + 1 \), where \( p \in [0,1] \) is the edge density. We then extend this result to sparse graphs by letting the edge density approach zero as the graph size grows. Importantly, the delay overhead due to uncertainty is still additive in Eq. (4). Nevertheless, the minimum delay the agent can afford has to be larger than that when the agent herself could design the network, as in Eq. (3).

4. General Graphs (Theorem 4) Finally, we turn to the general case, and provide upper and lower bounds on the minimal prediction risk that apply to any connected, undirected graph. The two bounds are tight up to a multiplicative factor, which depends on the level of uncertainty, \( \varepsilon \). We show that, for a graph with diameter \( d_G \),

\[
\frac{1}{2w+1} \leq Q(w) \leq \frac{2}{w(1-2\varepsilon) - d_G}.
\]

The above result hinges on a reduction argument that transforms an efficient agent strategy designed for the original deterministic Goal Prediction Game (Tsitsiklis and Xu (2018)) to our model with uncertainty. Setting \( \varepsilon = 0 \) in Eq. (5) recovers the bounds in Tsitsiklis and Xu (2018) for the deterministic model. On the positive end, Eq. (5) shows that, in general, the minimal prediction risk is not very sensitive to small perturbations in the level of uncertainty, \( \varepsilon \). On the other hand, the bounds in Eq. (5) are weaker than those in Eqs. (1), (3) and (4), in which the performance degradation due to uncertainty is additive instead of additive.
multiplicative; whether such a gap can be bridged can be an interesting direction for further research.

**Methodology.** The presence of uncertainty fundamentally changes the nature of decision making compared to that in a noiseless environment. In essence, this is because the presence of intrinsic uncertainty substantially restricts the type of randomizations that the agent is able to implement, by depriving her of the ability to pre-commit to a specific (randomly generated) trajectory before the game starts, a crucial feature to policy design in existing noiseless models (cf. Fanti et al. (2015), Tsitsiklis and Xu (2018), Cummings et al. (2016), Xu (2018)). Instead, the agent’s strategy must be adaptive and react to the actual realizations of the random shock, which significantly complicates the analysis. We develop new proof techniques and design principles to address these challenges raised by uncertainty, which depart substantially from the analysis in the noiseless model of Tsitsiklis and Xu (2018). At its core, we develop a family of strategies (Water-Filling) that “harvests” the intrinsic randomness and combines it with additional artificial randomization to prove Theorem 1. This concept of “uncertainty harvesting” is further developed in Theorems 2 and 3, where the graph topology is no longer complete.

1.3 Organization

The remainder of the paper is organized as follows. We formally describe our model in Section 2 and present our main results in 3. Section 4 discusses the related literature. Sections 6-8 are devoted to the proofs of our main results, with a proof overview in Section 5 that summarizes the key techniques. We conclude the paper with some discussion in Section 10.

2 Model: $\varepsilon$-Noisy Goal Prediction Game

We formally define the $\varepsilon$-Noisy Goal Prediction Game in this section, which builds upon the (noiseless) Goal Prediction Game in Tsitsiklis and Xu (2018) but with the additional element of uncertainty. Setting the noise level $\varepsilon = 0$ and horizon $K = \infty$ recovers the original game. We give a full description of the setup for completeness.

**Definition 1.** An $\varepsilon$-Noisy Goal Prediction Game is played between an agent and adversary, over discrete time $t \in \mathbb{N}$. An instance of the game consists of the following:

(a) an undirected graph $G = (\mathcal{V}, \mathcal{E})$, with $n$ vertices;
(b) an initial agent state $x_0 \in \mathcal{V}$;
(c) a time horizon, $K \in \mathbb{N}$;
(d) a goal vertex $D$ sampled uniformly at random from $\mathcal{V}$;
(e) a collection of i.i.d. Bernouilli random variables $\{B_t\}_{t=1}^{K}$ with success probability $1 - \varepsilon$;
(f) two mutually independent sequences of independent random variables $\mathcal{R}_A$ and $\mathcal{R}_D$.

In particular, the set of vertices $\mathcal{V}$ represents the states of the agent, and the set of edges $\mathcal{E}$ the allowed state transitions in each period. The random variables $\{B_t\}_{t=1}^{K}$ capture the intrinsic uncertainty in the system, where $B_t$ indicates whether the agent’s chosen action
at time $t$ will be fulfilled. Finally, $\mathcal{R}_A$ and $\mathcal{R}_D$ are independent random variables that can be used for the purpose of randomization, by the agent or the adversary, respectively.

**Agent Strategies and Trajectories.** An agent trajectory is a sequence of random variables $\{X_t, \Gamma_t\}_{t=1}^K$ where $X_t$ denotes the state of the agent at time $t$ and is a $\mathcal{V}$-valued random variable such that $(X_t, X_{t+1}) \in \mathcal{E}$, for all $t \geq 1$. Further, $\Gamma_t$ takes values in some arbitrary set and denotes any side information that is willfully provided to the adversary, by the agent at time $t$.

An agent strategy, $\psi$, is a mapping which sequentially generates the agent’s trajectory, by taking as input the graph $G$, initial state $x_0$, and the realized values of $D$ and $\mathcal{R}_A$ up to the present moment. In particular, at time $t$, the agent strategy $\psi$ generates action $a_t = \psi_t(H_t, \mathcal{R}_A, D)$ by choosing a vertex from among the neighboring vertices of the current state, $X_t$, using as input the history of the system up to time $t-1$: $H_t = \{(X_i, a_i, B_i)\}_{i=1}^{t-1} \cup \{X_t\}$. Then, the next state $X_{t+1}$ is equal to $a_t$ with probability $1 - \varepsilon$, and is sampled uniformly random from the neighboring vertices, otherwise:

$$P(X_{t+1} = v \mid X_t = x_t, a_t) = \begin{cases} 1 - \varepsilon + \frac{\varepsilon}{|\mathcal{E}(x_t)|}, & \text{if } v = a_t, \\ \frac{\varepsilon}{|\mathcal{E}(x_t)|}, & \text{if } v \notin \mathcal{E}(x_t) \setminus \{a_t\}, \\ 0, & \text{otherwise.} \end{cases}$$

where $\mathcal{E}(x_t)$ denotes the set of neighboring vertices of $x_t$, $\{v : (x_t, v) \in \mathcal{E}\}$.

Under any agent strategy $\psi$, the time at which the agent reaches the goal is referred to as the goal-hitting time, and is given by

$$T^\psi = \inf\{t \geq 1 : X_t = D\}. \quad (6)$$

Then, the delay of agent strategy $\psi$ is defined as $\mathbb{E}(T^\psi)$, where the expectation is taken with respect to the randomness in $D$, $\mathcal{R}_A$ and $\{B_t\}_{t=1}^K$.

If by the end of the time horizon $K$ the agent has not yet reached her goal, then at time $K+1$, she is automatically sent to $D$ and the adversary is declared the winner. Specifically, we set $X_{K+1} = D$ and assume $B_K = 1$. The horizon is a technical element of the model to ensure the integrability of $T^\psi$. For the most part, our analysis focuses on cases where $K$ is sufficiently large and has a negligible effect on the overall system dynamics.

**Adversary Strategies and Prediction Risk.** The adversary has access to $G$, $x_0$, $\mathcal{R}_D$ and the agent’s trajectory up to time $t$, i.e., $X_1, \ldots, X_t$, and $\Gamma_1, \ldots, \Gamma_t$. He does not observe the agent’s actual actions $\{a_t\}$. Based on this information, the adversary’s strategy, $\chi$, produces a decision variable $\hat{D}_t \in \mathcal{V} \cup \{0\}$. If $\hat{D}_t = 0$, we say that the adversary forgoes a prediction, and if $\hat{D}_t \in \mathcal{V}$, then a prediction is made. Accordingly, we let $U_{\psi, \chi}$ be the first time that the adversary makes a prediction, under a given pair $(\psi, \chi)$ of agent and adversary strategies, respectively:

$$U_{\psi, \chi} = \inf\{t \geq 1 : \hat{D}_t \in \mathcal{V}\}.$$

Thus, we can write $\hat{D}_{U_{\psi, \chi}}$ to represent the adversary’s prediction. Appendix A provides an illustrative instance of the $\varepsilon$-Noisy Goal Prediction Game.

We say that the adversary wins the game if his first prediction correctly matches the goal, and is made by the time the goal is reached by the agent. The adversary is again declared the winner if the agent fails to reach her goal vertex by the end of the horizon. We define prediction risk of a given pair of agent and adversary strategies $(\psi, \chi)$ as the
probability
\[
q(\psi, \chi) = \mathbb{P}(D_{U_{\psi, \chi}} = D \text{ and } U_{\psi, \chi} \leq T^\psi) + \mathbb{P}(T^\psi > K),
\]
with respect to the randomness in \(D, R_A, R_D, \text{ and } \{B_t\}_{t=1}^K\). Note that, for any fixed agent strategy \(\psi\), the probability that the agent loses the game due to not reaching the goal by the end of the horizon diminishes as the length of the horizon increases, i.e., \(\lim_{K \to \infty} \mathbb{P}(T^\psi > K) = 0\).

**Minimax Prediction Risk.** Given an agent strategy \(\psi\), we define the maximal prediction risk:
\[
q^*(\psi) = \sup_{\chi} q(\psi, \chi).
\]
For any given time budget \(w \in \mathbb{R}^+_0\), \(\Psi_w\) is defined to be the set of all agent strategies for which the delay is at most \(w\):
\[
\Psi_w = \{\psi : \mathbb{E}(T^\psi) \leq w\}.
\]
Finally, we define our main metric of interest, the minimax prediction risk:
\[
Q(w) = \inf_{\psi \in \Psi_w} q^*(\psi) = \inf_{\psi \in \Psi_w} \sup_{\chi} q(\psi, \chi).
\]
In words, \(Q(w)\) measures the least amount of prediction risk that any agent strategy will be able to guarantee subject to a delay of at most \(w\). Characterizing \(Q(w)\) will allow us to design efficient agent policies with limited resources, under uncertainty.

2.1 An Offline Prediction Variant
While a competitor in the secure logistics application might make online predictions by deciding whether to make a prediction at each period \(t\), in some applications such as secure messaging, the adversary might have the luxury to make an offline prediction after the whole trajectory has been realized. For instance, in the context of secure messaging an advertiser may collect data on network traffic over a communication platform for several months and analyze it later. Hence, the offline prediction variant may be very relevant for such an application. In fact, our results extend to the following offline model: the adversary needs to make a prediction at time \(K\), after observing the whole trajectory but without knowing whether the goal has been reached. On the other hand, recall that in the online model the adversary wins the game based on both the identity of his prediction and its timing.

We highlight that both the theorems and the strategies carry over to the offline setting. We first note that the lower bounds continue to hold. This is because in the online setting the agent is faced with a more powerful adversary, implying greater values of the prediction risk. Similarly, the analysis for the upper bounds will carry over since we will focus on agent strategies that reveal their trajectories at time 1, and prove the results against an offline adversary. Essentially, the strategies designed for the online setting extend to the offline variant and all of the theorems continue to hold.

2.2 Notation
Let \(f, g : \mathbb{N} \to \mathbb{R}\) be two functions. We use the following asymptotic notation: \(f(n) \ll g(n)\) if \(\lim_{n \to \infty} f(n)/g(n) = 0\); \(f(n) \sim g(n)\) if \(\lim_{n \to \infty} f(n)/g(n) = 1\); \(f(n) \leq g(n)\) if \(\limsup_{n \to \infty} f(n)/g(n) < \infty\); and \(f(n) \preceq g(n)\) if \(\limsup_{n \to \infty} f(n)/g(n) \leq 1\).
3 Main Results

We state the main results in this section. The proofs will be given in Sections 6 through 8, with an overview in Section 5.

The first theorem gives an asymptotically tight characterization of the prediction risk in complete graphs. We focus on the regime in which the number of vertices, $n$, is let to approach infinity.

**Theorem 1.** Fix noise level $\varepsilon \in [0, 1)$ and delay budget $w$ such that $w > \frac{1}{2} + \sqrt{\frac{\varepsilon}{1-\varepsilon}}$. Fix a sequence of horizons $\{K_n\}_{n \in \mathbb{N}}$ such that $1 \ll K_n \ll n$. Fix $n \in \mathbb{N}$, and let $G = (\mathcal{V}, \mathcal{E})$ be a complete graph with $n$ vertices. Then, the minimax prediction risk satisfies

$$\frac{1}{2w-1}\left(1-\alpha_\varepsilon(w)\right) - \delta_n \leq Q(w) \leq \frac{1}{2w-1}\left(1-\beta_\varepsilon(w)\right) + \delta_n,$$

where $\lim_{n \to \infty} \delta_n = 0$, and $\alpha_\varepsilon(w)$ and $\beta_\varepsilon(w)$ are defined by

$$\alpha_\varepsilon(w) = \frac{\varepsilon}{(2w-1)(1-\varepsilon)^2} > 0, \text{ and } \beta_\varepsilon(w) = \frac{\varepsilon^2}{2(w-1)^3(1-\varepsilon)^4} > 0.$$

Notably, for large complete graphs, Theorem 1 establishes that the delay overhead due to intrinsic uncertainty is only additive as a function of the noise level. Further, the overhead is always strictly positive regardless the size of the delay budget.

Theorems 2 and 3 examine non-complete graphs when the network topology can be designed by the agent and when it is given exogenously, respectively. Theorem 2 shows that we can design private networks with average degree as low as $\sqrt{n}$ while maintaining additive delay overhead.

**Theorem 2.** Fix noise level $\varepsilon \in [0, 1)$. Fix a sequence of horizons $\{K_n\}_{n \in \mathbb{N}}$ such that $1 \ll K_n \ll n$. Fix a degree sequence $\{\bar{p}_n\}_{n \in \mathbb{N}}$ such that $\bar{p}_n \ll n$ and $\bar{p}_n \gg \sqrt{n}$. Fix $n \in \mathbb{N}$. Then, there exists a family of graphs, $\mathcal{G}(n, \bar{p}_n)$, consisting of graphs with $n$ vertices and average degree at most $\bar{p}_n$. Fix $w$ such that $w > \left(\frac{1}{2} + \sqrt{\frac{\varepsilon}{1-\varepsilon}}\right)\frac{1}{\bar{p}_n}$. Then, the minimax prediction risk for this family satisfies

$$\frac{1}{2w+1} \leq Q(w) \leq \frac{1}{2w\bar{p}_n-1} + \delta_n,$$

where $\lim_{n \to \infty} \bar{p}_n = 1$ and $\lim_{n \to \infty} \delta_n = 0$, and $\varepsilon^c = \frac{\varepsilon}{(1-\varepsilon)^3}$.

In Section 7.3, we explain how to explicitly construct the graphs $\mathcal{G}(n, \bar{p}_n)$ that allow the agent to achieve an additive overhead.

The next theorem examines the case when the agent cannot design the network, which is in contrast to Theorem 2 where the agent could choose the network topology. Specifically, we design an agent strategy that can be implemented on any undirected graph and construct a family of graphs, $\mathcal{G}_n$, on which this strategy guarantees an additive delay overhead. We further establish that the family $\mathcal{G}_n$ is “large,” in the sense that an Erdős-Rényi random graph with $n$ vertices and edge probability $p$ belongs to $\mathcal{G}_n$ with high probability, as $n \to \infty$.

**Theorem 3.** Fix a noise level $\varepsilon \in [0, 1)$, edge density $p \in [0, 1]$ and delay budget $w > \frac{2}{p} \left(\frac{1}{2} + \sqrt{\frac{\varepsilon}{1-\varepsilon}}\right) + 1$. Fix a sequence of horizons $\{K_n\}_{n \in \mathbb{N}}$ such that $1 \ll K_n \ll n$. Then, there exists a sequence of families of graphs, $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$, such that the following are true.
(a) Suppose that $\tilde{G}$ is a graph drawn from the Erdős-Rényi random graph model with $n$ vertices and edge probability $p$. Then,

$$\mathbb{P}(\tilde{G} \in \mathcal{G}_n) \geq 1 - \theta_n,$$

where $\lim_{n \to \infty} \theta_n = 0$.

(b) Fix $n \in \mathbb{N}$, and suppose that $G \in \mathcal{G}_n$. Then, the minimax prediction risk satisfies

$$\frac{1}{2w + 1} \leq Q(w) \leq \frac{1}{w - \lambda(p)} + \delta_n,$$

where $\delta_n$ is a constant that does not depend on $G$ and $\lim_{n \to \infty} \delta_n = 0$, and $\lambda(p) = 1 + \frac{1}{p} + \frac{\varepsilon}{(1 - \varepsilon)^2 p}$.

Theorem 3 demonstrates that additive overhead on delay can still be achieved on a “typical” non-complete graph, generated by a random graph model. In Appendix G, we generalize the analysis of Theorem 3 to families of sparse graphs where the edge density, $p$, can decrease as the graph size, $n$, increases. There we show that a similar performance can be attained for “typical” sparse non-complete graphs.

Lastly, for any connected undirected graph, we design a policy that can be implemented under any intrinsic noise level, which leads to the following result. We denote the diameter of a graph $G$ by $d_G = \max_{v,v' \in V} d(v, v')$, where $d(v, v')$ is the number of edges in the shortest path from $v$ to $v'$.

**Theorem 4.** Fix noise level $\varepsilon \in [0, 1)$. Let $G = (\mathcal{V}, \mathcal{E})$ be a connected undirected graph. Fix $w$ such that $w > \frac{d_G}{1 - 2\varepsilon}$. Then, the minimax prediction risk satisfies

$$\frac{1}{2w + 1} \leq Q(w) \leq \frac{2}{w(1 - 2\varepsilon) - d_G},$$

(10)

First, setting $\varepsilon = 0$ recovers the result of Tsitsiklis and Xu (2018). Second, the overhead due to the intrinsic uncertainty degrades multiplicatively as the noise level increases. Furthermore, the overhead is small for low noise levels, implying that the system performance is not very sensitive to small perturbations.

## 4 Related Work

Our work is inspired by, and builds upon, the Goal Prediction Game proposed by Tsitsiklis and Xu (2018), which shows that, in a model with perfect control, the agent’s prediction risk is inversely proportional to her delay and this property holds essentially universally regardless of the detailed graph topology. Our formulation overcomes a crucial limitation of the model in Tsitsiklis and Xu (2018) by allowing the agent’s actual trajectory to be perturbed by noise and uncertainty. As discussed in the Introduction, this generalization provides a more powerful framework to capture a wider range of applications in which an agent is faced with a highly stochastic environment. On the methodological front, the incorporation of uncertainty and noise significantly complicates system dynamics. As a result, we will develop new algorithms and proof techniques that are distinct from those in Tsitsiklis and Xu (2018).
At a higher level, our work relates to a growing literature on privacy preserving mechanisms in dynamic decision-making, including those in operations research (cf. Cummings et al. (2016), Tsitsiklis and Xu (2018)), game theory (cf. Blum et al. (2015), Gradwohl and Smorodinsky (2017), Augenblick and Bodoh-Creed (2018)), computer science (cf. Fanti et al. (2015), Lindell and Pinkas (2009)) and learning theory (cf. Calmon et al. (2015), Shokri and Shmatikov (2015), Tossou and Dimitrakakis (2016), Xu (2018), Tsitsiklis et al. (2018)). However, in almost all cases, the key assumption is that the decision maker has full control over her actions, and the efficacy of a mechanism critically depends on the nature of randomness injected. For instance, Fanti et al. (2015) introduce a messaging protocol with a precise spreading probability at each time step to ensure the obfuscation of the rumor source, whereas Tsitsiklis and Xu (2018) rely on carefully constructing a set of paths, from which the agent will choose one to follow beforehand. In contrast, we assume an agent without full control over the consequence of her actions and aim to understand the impact of such uncertainty.

More broadly, our model is connected to the body of literature on privacy algorithms (cf. Chaudhuri et al. (2011), Jain et al. (2012), Gupta et al. (2012), Tossou and Dimitrakakis (2016)), among which differential privacy (Dwork et al. (2014)) is a well-known paradigm. Among many divergences, our framework differs from this literature first and foremost by focusing on goal-oriented privacy as opposed to a universal privacy criterion. That is, we define privacy with respect to a specific statistical inference task that the adversary has, i.e., predicting the goal of the agent. Differential privacy, on the other hand, imposes a significantly more stringent requirement that the mechanism be private with respect to any inference task. As such, our formulation, like other goal-oriented privacy models (cf. Calmon et al. (2015), Fanti et al. (2015), Issa et al. (2016), Radaelli et al. (2018), Xu (2018)) is able to deliver stronger performance guarantees for a given level of desirable privacy, whereas universal privacy provides stronger privacy guarantees but at the expense of a higher efficiency loss.

5 Proof Overview

The next three sections are devoted to the proofs of the main theorems. Before we delve into the details, it is instructive to understand why the incorporation of uncertainty makes the design of private agent strategies more difficult. In the noiseless setting (ε = 0), the family of near-optimal agent strategies proposed by Tsitsiklis and Xu (2018) can be roughly summarized as follows. The agent’s strategy can be viewed as a random mapping S that takes as input the goal vertex D, and outputs a path that the agent will follow. The mapping S is designed in such a way that: (a) the length of S(D) is at most w, ensuring a small delay, and (b) S(D) does not significantly reveal D, ensuring privacy. Tsitsiklis and Xu (2018) shows how to construct S for general graphs, by first creating a family of paths, S, that covers the graph in a uniform manner, and then generating S(D) by sampling uniformly at random from those paths in S that contain D. Importantly, in the noiseless setting, because the agent has full control over her future trajectory, the path S(D) only depends on D and can thus be generated at the beginning in a non-adaptive manner.

Unfortunately, the above framework breaks down when uncertainty is present. The presence of noise at each step means that the agent can no longer implement a specific path at will. She is confronted with two design options. She may still “emulate” a policy designed for the noiseless model by trying to follow a pre-sampled path, S(D). However, it will now
take multiple steps for the agent to traverse a single step in \( \mathcal{S}(D) \). In the proof of Theorem 4, we show that, while this is indeed a viable option that works on general graphs, the agent will have to suffer a multiplicative delay increase.

Alternatively, the agent may fundamentally change her approach by using an adaptive strategy. Instead of trying to follow a pre-determined path, her actions can now depend on past realizations of the intrinsic uncertainty. Theorems 1, 2 and 3 demonstrate that this approach of adaptively “harvesting” intrinsic uncertainty is far superior, and allows the agent to achieve optimal to near-optimal performance in a variety of network topologies.

We will present the proofs of the main theorems with a different sequence as the one in which they were originally stated in Section 3. We will start by showing the results for general graphs, Theorem 4, followed by Theorems 1 (complete graphs), 2 (private network design) and 3 (non-complete graphs). The intention here is two-fold. First, the proof of Theorem 4 is the simplest, as it follows from a fairly natural extension of strategies designed for the original noiseless model. More importantly, as alluded to earlier, the proof of Theorem 4 will help illustrate a key difficulty in dealing with intrinsic uncertainty, where the agent will no longer be able to commit to a pre-determined trajectory. The results in the subsequent sections will address this challenge, and achieve additive delay overhead with more sophisticated, adaptive strategies.

6 General Graphs: Proof of Theorem 4

We will prove the upper bound in Theorem 4 by a reduction argument that adapts an efficient agent strategy designed for noiseless Goal Prediction Game (\( \varepsilon = 0 \)) to our model with uncertainty. For the remainder of this proof, we will use the term “noiseless” to signify the case where \( \varepsilon = 0 \), and “noisy” for when \( \varepsilon > 0 \). We will first analyze the case where the time horizon \( K \) is assumed to be infinite and later extend the result to \( K < \infty \).

We will choose as our noiseless baseline the so-called trajectory revealing agent strategies. Note that in a noiseless model, the agent will be able to pick a trajectory in advance and perfectly execute it. We say a noiseless agent strategy is trajectory revealing if it announces to the adversary at \( t = 1 \) the entire future trajectory. While doing so might sound counter-intuitive (why would the agent reveal her actions in advance?!), a trajectory revealing strategy has the benefit that, once deployed, the adversary has incentive to simply make a prediction at \( t = 1 \), since future agent actions will not reveal any additional information. Furthermore, it is shown in Tsitsiklis and Xu (2018) that there exists a family of trajectory revealing strategies (segment-based strategies) that achieve a near-optimal delay-privacy tradeoff. In particular, for any \( w \), with \( d_G < w \leq n \), there exists a trajectory revealing strategy, \( \psi_0 \), with delay at most \( w \), and prediction risk at most \( \frac{2}{w-d_G} \). This is possible, since the trajectory is randomly generated such that even when it is revealed, it remains unclear which vertex along the trajectory is the goal.

Fix a delay target of \( w' \), and a corresponding noiseless trajectory revealing strategy \( \psi_0 \). We now show how to convert the noiseless agent strategy \( \psi_0 \) to one that works in the noisy setting. To do so, we first introduce the concept of a greedy shortest path. For two vertices \( v, v' \in \mathcal{V} \), a greedy shortest path (GSP) routine from \( v \) to \( v' \) is an agent strategy under which the agent chooses the next vertex on the shortest path from her current state to \( v' \) in each period. We will denote by \( L(v, v') \) the expected number of steps to travel from \( v \) to \( v' \) under a GSP routine.

The following result will be used shortly, which provides an upper bound on \( L(v, v') \).
The proof is based on a coupling argument to a random walk on $G$, and is given in Appendix E.1.

**Proposition 1.** Let $G = (V, E)$ be a connected undirected graph and fix $v, v' \in V$. Then, the length of the greedy shortest path between $v$ and $v'$ satisfies

$$L(v, v') \leq \frac{d(v, v')}{1 - 2\varepsilon}. \tag{11}$$

We are now ready to describe the agent strategy that achieves the upper bound. At time $t = 1$, the agent first generates a trajectory, $\{X_t\}_{t \in \mathbb{N}}$ using the noiseless trajectory revealing strategy $\psi_0$, described above, and announces this trajectory to the adversary via $\Gamma_1$. Subsequently, the agent traverses the entirety of $\{X_t\}_{t \in \mathbb{N}}$ sequentially by using GSP routines between two consecutive vertices. That is, the agent will employ a GSP routine to go from $X_1$ until she reaches $X_2$, and from there employ a GSP routine again to reach $X_3$, and so on, until the entirety of $\{X_t\}$ has been traversed. Furthermore, during this process, the agent will explicitly inform the adversary whether she is currently on an element of $\{X_t\}$ versus performing a GSP routine. Intuitively, this strategy “forces” pre-commitment by implementing a pre-determined trajectory using GSP routines. We will thus refer to it as the *Stochastic Pre-Commit Strategy*, denoted by $\psi_{SPC}$.

**Proposition 2.** Let $\psi_0$ be a trajectory revealing strategy in the noiseless setting and $\psi_{SPC}$ be the corresponding Stochastic Pre-Commit Strategy. Then, $q^*(\psi_{SPC}) = q^*(\psi_0)$.

Using Proposition 2, we obtain that the maximal prediction risk of the Stochastic Pre-Commit Strategy is at most $w' - d_G$ if the delay budget for $\psi_0$ is $w'$ in the noiseless setting. The delay of Stochastic Pre-Commit is, however, enlarged compared to $\psi_0$. Note that by construction, the delay associated with the noiseless trajectory $\{X_t\}_{t \in \mathbb{N}}$ is at most $w'$. Because traversing each pair of adjacent vertices $(X_t, X_{t+1})$ now consumes an expected number of steps, by Proposition 1, we conclude that the delay of $\psi_{SPC}$ satisfies

$$E(T_{SPC}) \leq \frac{1}{1 - 2\varepsilon}w'. \tag{12}$$

Combining the above two observations, and by substituting $w'$ with $w(1 - 2\varepsilon)$, we have thus identified a Stochastic Pre-Commit strategy with a delay that is at most $w$, and a prediction risk that is at most $\frac{2}{w(1 - 2\varepsilon) - d_G}$. This completes the proof of the upper bound in Theorem 4, with $K = \infty$.

We now turn to the case where $K < \infty$. When the horizon is finite, the agent can continue using the Stochastic Pre-Commit strategy with delay budget $w$. The truncation at time $K$ will cause the expected delay to be smaller and the same strategy will continue satisfying the delay budget. Regarding the prediction risk, the probability of not reaching the goal before $K$ is simply $P(T > K)$. Then, since $\lim_{K \to \infty} P(T > K) = 0$ holds, the difference will become negligible in the regime of interest.

For the lower bound in Theorem 4, we note that the addition of noise makes the strategy space of the agent more restrictive than the noiseless setting ($\varepsilon = 0$), since in the latter case
the agent can also simulate the noise in her state transition if necessary. As a result, the prediction risk lower bound in Tsitsiklis and Xu (2018) for the deterministic setting still holds in our case. The lower bound thus follows from the lower bound in Corollary 1 of Tsitsiklis and Xu (2018). This concludes the proof of Theorem 4.

7 Water-Filling Strategy on Complete Graphs

We prove Theorem 1 in this section, which will serve as the foundation of the analysis for all non-complete graphs in subsequent sections. As we saw in the preceding section, it is possible to adapt an effective agent strategy from the noiseless model ($\varepsilon = 0$) to our noisy setting ($\varepsilon > 0$) by traversing a pre-determined trajectory using greedy shortest path routines. Yet, doing so means that the agent incurs a delay on the order of $\frac{1}{\varepsilon}$ on every step along the trajectory, which ultimately leads to the multiplicative increase in delay overhead in Theorem 4.

In this section, we will take a different approach: instead of viewing the intrinsic uncertainty as necessary evil to be coped with, we may also harvest such randomness and combine it with the artificial randomization already used by the agent strategy (even in a noiseless setting). This further implies that we will have to let go the idea of a forced pre-commitment (such as in the Stochastic Pre-Commit Strategy) and adjust the agent’s actions in an adaptive manner. This will be achieved by the Water-Filling Strategy.

The analysis in this section, and in particular the proof of the upper bound and the construction of the Water-Filling Strategy, will form the basis of the proofs in the subsequent Sections 7.3 and 8, where we extend the analysis to study the issue of network design with non-complete graphs and to study random graphs.

7.1 Upper Bound in Theorem 1

We prove the upper bound in Theorem 1 in this subsection. The proof will rely on two main ideas which we highlight below. The first helps us characterize the delay of an agent strategy, and the second targets the analysis of the prediction risk.

(1) Intentional Goal-Hitting Time. We will make repeated use of the concept of intentional goal-hitting time, defined as the first time the goal is reached by the agent intentionally (i.e., when it coincides with her action), as opposed to being the result of a random shock:

$$T_{IH}^\psi = \inf\{t \geq 1 : X_t = D \text{ and } B_{t-1} = 1\}. \quad (13)$$

The usefulness of $T_{IH}^\psi$ stems from the fact that, in a model with uncertainty, it is significantly easier to characterize when the agent intends to reach the goal (something she can control), versus when she actually hits it. As such, the intentional goal-hitting time provides a natural upper bound on the actual goal-hitting time, and is more amenable to analysis.

(2) Trajectory-Goal Independence. Deriving strong upper bounds on the prediction risk requires the analysis to have a precise control over how much information the adversary can learn by observing the agent’s past trajectory.\(^3\) In a noiseless model with perfect control,

\(^2\)Note that $T_{IH}^\psi \leq K + 1$: if the agent has not yet reached her goal by the end of the horizon, then we have $T_{IH} = K + 1$ since we set $X_{K+1} = D$ and $B_K = 1$.

\(^3\)As an example, Appendix B shows an instance where an agent strategy may inadvertently reveal the identity of the goal, even if its goal-hitting time has a nearly-uniform distribution.
this is more easily achievable because the agent strategy is able to pre-commit to a specific trajectory. Unfortunately, this is no longer the case when uncertainty is present.

In the proof that follows, we will control the information leakage by carefully creating an independence between the trajectory and the agent’s goal. (Luckily, in the case of a large complete graph, or graphs we construct in Sections 7.3 and 8, such independence can be implemented with no, or relatively little, delay overhead.) We will employ a number of techniques to achieve this objective, including a path counting argument, the notion of the intentional goal-hitting time, as well as the design of the Water-Filling Strategy itself.

We next introduce the main strategy of this section: the Water-Filling Strategy, which aims to achieve a small expected intentional goal-hitting time while ensuring the probability of being the intentional hitting time is upper-bounded by some target risk level, \( \bar{q} \), for any period \( t \).

To simplify the analysis, we will design the Water-Filling Strategy such that at every period the agent chooses one of two “meta” actions. We will denote by \( \bar{A} \) the set of meta actions and by \( \bar{a}_t \in \bar{A} \) the meta action taken at time \( t \). The two meta actions are defined as follows.

1. **Random** (\( \bar{a}^R \)). Under \( \bar{a}^R \), the agent samples a vertex uniformly at random from all vertices as her action.

2. **Goal-Attempt** (\( \bar{a}^G \)). Under \( \bar{a}^G \), the agent chooses the goal vertex \( D \) as her action.

Crucially, these meta actions are intended to highlight when the agent intentionally reaches the goal, versus reaching the goal by chance.

Recall that if \( B_t = 1 \), then the agent’s action in period \( t \) has been successful whereas she has been sent to a random vertex, otherwise. With this notation, we define the sequence \( \{ \bar{F}_t \}_{t \in \mathbb{N}} \) of indicators as follows: if the agent’s meta action at time \( t \) is Goal-Attempt (\( \bar{a}_t = \bar{a}^G \)) and the goal is in fact reached successfully, i.e., \( X_{t+1} = D \) and \( B_t = 1 \), then \( \bar{F}_{t+1} = 1 \). Otherwise, \( \bar{F}_{t+1} = 0 \). Using this sequence, the intentional goal-hitting time, that is, the first time the agent chooses the meta action Goal-Attempt (\( \bar{a}_t = \bar{a}^G \)) and actually succeeds, can be written as \( T_{IH} = \inf \{ t \geq 1 : \bar{F}_t = 1 \} \).

We now describe the main agent strategy in this section, dubbed the Water-Filling Strategy.\(^4\)

**Definition 2 (Water-Filling Strategy).** Fix a target risk level \( \bar{q} \in [0, 1] \). Let

\[
\bar{t}^* = \left\lfloor \frac{1}{\bar{q}} - \frac{\varepsilon}{1 - \varepsilon} \right\rfloor.
\]

Define

\[
p_t = \begin{cases} 
\frac{\bar{q}}{1 - \bar{q}} (1 - t\bar{q})^{-1}, & \text{if } 0 \leq t < \bar{t}^* - 1, \\
1, & \text{otherwise}.
\end{cases}
\]

The Water-Filling Strategy, \( \psi_{\bar{q}}^{wf} \), on a complete graph \( G = (V, E) \) with \( n \) vertices is defined recursively as follows.

\(^4\)To simplify notation and to avoid floor and ceiling functions, we assume the parameters are such that \( \bar{t}^* \) is an integer.
(1) If $\max_{1 \leq s \leq t} \tilde{F}_s = 0$, i.e., a successful goal attempt has not occurred, then the agent will choose the meta action Goal-Attempt ($\tilde{a}_s = \tilde{a}^G$) with probability $p_t$, and Random ($\tilde{a}_s = \tilde{a}^R$), otherwise.

(2) If $\max_{1 \leq s \leq t} \tilde{F}_s = 1$, i.e., a successful goal attempt has already occurred, then the agent will choose the meta action Random ($\tilde{a}_s = \tilde{a}^R$).

**Remark.** (Interpretation of the Water-Filling Strategy) Intuitively, in order to achieve a desirable prediction risk, the agent must construct a strategy under which the distribution of the goal-hitting time is nearly uniform, for otherwise any abnormally high peak in the probability mass function of the goal-hitting time would compromise her privacy. However, since nature is injecting noise into the system in every step, implementing the uniform distribution exactly becomes infeasible. To address this problem, the Water-Filling Strategy mimics the uniform distribution in a recursive manner until it is no longer possible to do so, while in the mean time ensuring that the trajectory, $\{X_t\}_{t \in \mathbb{N}}$, and the intentional goal-hitting time, $T_{\text{IH}}^{\text{wf}}$, remain mutually independent.

Specifically, the agent begins by choosing the meta action Goal-Attempt with probability $p_t$ such that the probability that intentional goal-hitting time is $t + 1$ is exactly equal to $\tilde{q}$, i.e., $P(T_{\text{IH}}^{\text{wf}} = t + 1) = \tilde{q}$. Starting at time $t^* - 1$, however, she can no longer achieve $P(T_{\text{IH}}^{\text{wf}} = t + 1) = \tilde{q}$ because to do so she would need to set $p_t > 1$, which is impossible. Consequently, she starts attempting the goal with probability 1, resulting in $P(T_{\text{IH}}^{\text{wf}} = t) < \tilde{q}$ for $t \geq t^*$. A graphical illustration of the distribution of $T_{\text{IH}}^{\text{wf}}$ is given in Figure 1.

![Figure 1: Distribution of $T_{\text{IH}}^{\text{wf}}$ under the Water-Filling Strategy, with $\tilde{q} = 0.21$ and $\varepsilon = 0.5$. The vertical line depicts the threshold $t^*$.

In a nutshell, as shown in Figure 1, the strategy aims to set $P(T_{\text{IH}}^{\text{wf}} = t)$ to be equal to the upper bound $\tilde{q}$ for as long as possible in a greedy manner, and as a result, it ensures that $E(T_{\text{IH}}^{\text{wf}})$ is minimal subject to the constraint on prediction risk. Moreover, the strategy uses the identity of $D$ only to compute the indicators $\tilde{F}_t$ and is completely symmetric across all other actions. This structure of the strategy guarantees the independence between the trajectory and $T_{\text{IH}}^{\text{wf}}$, so that the adversary cannot effectively learn $T_{\text{IH}}^{\text{wf}}$ from the observed trajectory. These factors together will allow us to establish the desirable delay vs. prediction risk tradeoff in Theorem 1.

We now proceed to the analysis of the agent’s prediction risk and delay under the Water-Filling Strategy. Proposition 3 derives the expected intentional goal-hitting time under $\psi^{\text{wf}}$. 
and is proven in Appendix E.3.\textsuperscript{5} Crucially, note that the noise level $\varepsilon$ has an additive effect on the delay, in contrast to the per step overhead we had for general graphs.

**Proposition 3.** Fix graph size $n \in \mathbb{N}$, horizon $K \in \mathbb{N}$, and target risk level $\bar{q} \in [0, 1]$. Let $G = (\mathcal{V}, \mathcal{E})$ be a complete graph with $|\mathcal{V}| = n$. Under the Water-Filling Strategy, $\psi_{w}^{\text{wf}}$, the expected goal-hitting and intentional goal-hitting times satisfy

$$
E(T_{\text{wf}}) \leq E(T_{\text{IH}}) \leq \frac{1}{2\bar{q}} + \frac{1}{2} + \frac{\bar{q}\varepsilon}{2(1 - \varepsilon)^2}.
$$


Now, we turn to the prediction risk of $\psi_{w}^{\text{wf}}$ and state that the maximal prediction risk is asymptotically upper-bounded by the target risk level. The proof of Proposition 4 is given in Appendix E.4 and relies on two key technical results. These technical results and the insight derived from the proof of Proposition 4 will further be used in Section 8 to prove Theorems 3 and 5, though incorporating additional complexity and richness.

**Proposition 4.** Fix a sequence of horizons $\{K_n\}_{n \in \mathbb{N}}$ such that $1 \ll K_n \ll n$. Fix graph size $n \in \mathbb{N}$, and target risk level $\bar{q} \in [0, 1]$. Let $G = (\mathcal{V}, \mathcal{E})$ be a complete graph with $|\mathcal{V}| = n$. Under the Water-Filling Strategy, $\psi_{w}^{\text{wf}}$, the maximal prediction risk satisfies

$$
q^*(\psi_{w}^{\text{wf}}) \leq \bar{q} + \delta'_n,
$$

where $\delta'_n = \frac{w}{K_n} + 1 - \left(1 - \frac{1}{n}\right)K_n$, and $\lim_{n \to \infty} \delta'_n = 0$.

### 7.1.1 Proof of the Upper Bound

Combining the results proven so far, we conclude the proof of the upper bound in Theorem 1, in which the cost of noise is only additive as opposed to being multiplicative as before. To this end, fix a delay budget $w$ and consider the class of strategies, $\Phi_{w}$, under which the expected intentional goal-hitting time is at most $w$, i.e., $\Phi_{w} = \{\psi : E(T_{\text{IH}}^{\psi}) \leq w\}$.

Define $\bar{q}(w)$ to be the minimum target risk level $\bar{q}$ that can be attained by a Water-Filling Strategy with expected intentional goal-hitting time at most $w$, that is, $\bar{q}(w) = \inf\{\bar{q} : \exists \psi_{w}^{\text{wf}} \in \Phi_{w}\}$.

Now, we explicitly calculate $\bar{q}(w)$. Suppose $w > \frac{1}{2} + \frac{\sqrt{2}}{1 - \varepsilon}$. Using Proposition 3 and solving for $E(T_{\text{IH}}^{\psi}) = w$ subject to $\bar{q} \in [0, 1]$, we obtain

$$
\bar{q}(w) = \frac{1}{(w - \frac{1}{2}) \left(1 + \frac{1}{\sqrt{1 - \frac{\varepsilon^2}{(w - \frac{1}{2})^2(1 - \varepsilon)^2}}}\right)}.
$$

Let $x = \frac{\varepsilon}{(w - \frac{1}{2})^2(1 - \varepsilon)^2}$. By the Taylor expansion of $\sqrt{1 - x}$ for $x < 1$, we have that

$$
\bar{q}(w) \leq \frac{1}{2w - 1 - \alpha^\varepsilon(w) - \beta^\varepsilon(w)},
$$

where $\alpha^\varepsilon(w) = \frac{\varepsilon^2}{(2w - 1)^2(1 - \varepsilon)^2}$ and $\beta^\varepsilon(w) = \frac{\varepsilon^2}{2(w - \frac{1}{2})^2(1 - \varepsilon)^2}$.

\textsuperscript{5}In the remainder, we may suppress the argument $\bar{q}$ from $\psi_{w}^{\text{wf}}$ for notational convenience, when the context is clear.
Since setting $q = \bar{q}(w)$ implies $\mathbb{E}(T_{\text{WF}}) \leq \mathbb{E}(T_{\text{IH}}) \leq w$, we obtain
\[ Q(w) \leq \frac{1}{2w - 1 - \alpha(w) - \beta(w)} + \delta_n', \] (16)
using Proposition 4 and Eq. (15).

This concludes the proof of the upper bound in Theorem 1. Observe that the overhead due to noise is only additive.

7.2 Lower Bound in Theorem 1

So far, we have demonstrated that it is possible for the agent to harvest uncertainty to achieve an additive overhead in delay. In this section, we will focus on the lower bound and show that the agent always has to bear this strictly positive overhead due to the inherent uncertainty, even if the delay budget $w$ is large. That is, the effect of even a mild amount of uncertainty can never be fully naturalized.

To prove the lower bound in Theorem 1, we follow three main steps. We begin by showing that when the number of vertices is sufficiently large, $\mathbb{E}(T_{\psi})$ and $\mathbb{E}(T_{\text{IH}}^\psi)$ coincide. Second, for any strategy $\psi$, we define its maximal time-only prediction risk, $\hat{q}_\psi$, as its maximal prediction risk when the adversary strategy only uses the distribution of $T_{\text{IH}}^\psi$ to make a prediction, and not the actual trajectory. Consequently, we conclude that the maximal prediction risk is asymptotically lower bounded by $\hat{q}_\psi$. Third, we establish that among all agent strategies that can be defined on a complete graph $G$ with $n$ vertices, the strategy $\psi_{\text{WF}}^{\bar{q}}$ minimizes $\mathbb{E}(T_{\text{IH}}^\psi)$ subject to the constraint that $\hat{q}_\psi$ is at most $\bar{q}$. Thus, for any delay budget, we can assert that the minimum target risk level that can be attained by a Water-Filling Strategy with this budget is a lower bound for the maximal time-only prediction risk for any strategy with that budget. Note that we can also conclude the Water-Filling Strategy is optimal. Precisely, given any target risk level it minimizes the delay and given any delay budget it achieves minimal prediction risk.

Lemma 1 states that for sufficiently large complete graphs, $\mathbb{E}(T_{\psi})$ and $\mathbb{E}(T_{\text{IH}}^\psi)$ are equivalent so that we can use $\mathbb{E}(T_{\text{IH}}^\psi)$ as a proxy for $\mathbb{E}(T_{\psi})$. The proof is given in Appendix E.5.

**Lemma 1.** Fix a sequence of horizons $\{K_n\}_{n \in \mathbb{N}}$ such that $1 \ll K_n \ll n$. Fix graph size $n \in \mathbb{N}$. Let $G = (V,E)$ be a complete graph with $|V| = n$. Let $\psi$ be an agent strategy on $G$. Under $\psi$, the expected goal-hitting and intentional goal-hitting times satisfy
\[ \mathbb{E}(T_{\psi}^\phi) \leq \mathbb{E}(T_{\psi}^\phi) + \sigma_n, \]
where $\sigma_n = \left(1 - \left(1 - \frac{1}{n}\right)^{K_n}\right)K_n(K_n+1)/2$, and $\lim_{n \to \infty} \sigma_n = 0$.

Note that Lemma 1 allows us to state $\mathbb{E}(T_{\psi}^\phi) \leq \mathbb{E}(T_{\text{IH}}^\psi) \leq \mathbb{E}(T_{\psi}^\phi) + \sigma_n$, since the first inequality always holds. Thus, for sufficiently large graphs $\mathbb{E}(T_{\text{IH}}^\psi)$ is a good approximation for $\mathbb{E}(T_{\psi}^\phi)$.

Now, we analyze the maximal prediction risk. For any strategy $\psi$, define its maximal time-only prediction risk,
\[ \hat{q}_\psi = \max_{t \in \mathbb{N}} \mathbb{P}(T_{\text{IH}}^\psi = t). \]
The next lemma shows that the maximal prediction risk under any strategy $\psi$ is lower-bounded by its maximal time-only prediction risk, $\tilde{q}_\psi$, as the number of vertices increases. For sufficiently large graphs, this allows us to use the maximal time-only prediction risk, $\tilde{q}_\psi$ as a lower bound for $q^*(\psi)$. The proof is given in Appendix E.6.

**Lemma 2.** Fix a sequence of horizons $\{K_n\}_{n \in \mathbb{N}}$ such that $1 \ll K_n \ll n$. Fix graph size $n \in \mathbb{N}$. Let $G = (\mathcal{V}, \mathcal{E})$ be a complete graph with $|\mathcal{V}| = n$. Let $\psi$ be an agent strategy on $G$. Under $\psi$, the maximal prediction risk satisfies

$$q^*(\psi) \geq (1 - \bar{\delta}_n)\tilde{q}_\psi,$$

where $\bar{\delta}_n = 1 - (1 - \frac{1}{n})^{K_n}$, and $\lim_{n \to \infty} \bar{\delta}_n = 0$.

We now use Lemmas 1 and 2 to derive the lower bound. Recall that $\Phi_w$ denotes the class of strategies under which the expected intentional goal-hitting time is at most $w$. Define $\tilde{q}(w)$ to be the minimum $\tilde{q}_\psi$ value attained by some strategy in this class, where $\tilde{q}_\psi = \max_{t \in \mathbb{N}} \mathbb{P}(T^\psi_{IH} = t)$ as before. That is, let

$$\tilde{q}(w) = \inf_{\psi \in \Phi_w} \tilde{q}_\psi.$$ 

For any $\psi$ such that $\mathbb{E}(T^\psi_{IH}) \leq w$, Lemma 1 implies $\mathbb{E}(T^\psi_{IH}) \leq w + \sigma_n$ and consequently, $\Psi_w \subseteq \Phi_{w+\sigma_n}$. From this observation, for any $\psi$, we can write

$$\inf_{\psi \in \Psi_w} \tilde{q}_\psi \geq \inf_{\psi \in \Phi_{w+\sigma_n}} \tilde{q}_\psi = \tilde{q}(w + \sigma_n). \quad (17)$$

Then, we obtain

$$Q(w) = \inf_{\psi \in \Psi_w} q^*(\psi) \geq \inf_{\psi \in \Phi_{w+\sigma_n}} (1 - \bar{\delta}_n)\tilde{q}_\psi \geq (1 - \bar{\delta}_n) \inf_{\psi \in \Psi_w} \tilde{q}_\psi \geq (1 - \bar{\delta}_n)\tilde{q}(w + \sigma_n), \quad (18)$$

where (a), (b) and (d) follow from the definition of $Q(w)$, Lemma 2, and Eq. (17), respectively. Finally, (c) is due to the observation that $1 - \bar{\delta}_n \geq 0$ and is independent of $\psi$.

To conclude the lower bound from Eq. (18), we need the following result on the optimality of $\psi^{WF}$. Proposition 5 establishes that the Water-Filling Strategy actually achieves minimal expected intentional goal-hitting time subject to the maximal time-only prediction risk being upper-bounded by the target risk level. The proof can be found in Appendix E.7.

**Proposition 5.** Fix graph size $n \in \mathbb{N}$ and $\bar{q} \in [0, 1]$. Let $G = (\mathcal{V}, \mathcal{E})$ be a complete graph with $|\mathcal{V}| = n$. The Water-Filling Strategy with target risk level $\bar{q}$ solves the optimization problem,

$$\min_{\psi \in \Psi} \mathbb{E}(T^\psi_{IH}) \text{ s.t. } \max_{t \in \mathbb{N}} \mathbb{P}(T^\psi_{IH} = t) \leq \bar{q},$$

where $\Psi$ denotes the set of all agent strategies defined on $G$.

Recall that $\tilde{q}(w)$ is the minimum target risk level $\tilde{q}$ that can be attained by a Water-Filling Strategy with expected intentional goal-hitting time at most $w$. Then, by Lemma 1, we have that

$$\tilde{q}(w) = \inf\{\tilde{q} : \exists \psi^{WF}_\tilde{q} \in \Phi_w\} \leq \tilde{q}(w + \sigma_n) \leq \inf\{\tilde{q} : \exists \psi^{WF}_\tilde{q} \in \Phi_{w+\sigma_n}\},$$

20
since \( \Psi_{w-s_n} \subseteq \Phi_{w} \). In addition, by Proposition 5, if under any strategy \( \psi \) we have \( \mathbb{E}(T_{IH}^{\psi}) < \mathbb{E}(T_{IH}^{\psi_w}) \), then it must be the case that \( \tilde{q}_{\psi} > \tilde{q}_{\psi_w} \). Therefore, using Proposition 5 and Eq. (18), we obtain

\[
Q(w) \geq (1 - \check{\delta}_n)\tilde{q}(w + \sigma_n).
\]

To complete the lower bound, we recall the derivation of \( \tilde{q}(w) \) and again use the Taylor expansion of \( \sqrt{1 - x} \) for \( x < 1 \). Consequently, letting \( \alpha^\varepsilon(w) = \frac{\varepsilon}{(2w-1)(1-x)^2} \), we obtain

\[
\frac{1}{2w-1-\alpha^\varepsilon(w)} \leq \tilde{q}(w).
\]

Then, combining Eqs. (19) and (20) gives

\[
Q(w) \geq \frac{1 - \check{\delta}_n}{2w-1-\alpha^\varepsilon(w) + \sigma_n} \geq \frac{1 - \check{\delta}_n}{2w-1-\alpha^\varepsilon(w) + \sigma'_n} \geq \frac{1}{2w-1-\alpha^\varepsilon(w)} - \check{\delta}_n,
\]

where (a) and (b) are obtained by the Taylor expansion of \( \alpha^\varepsilon(w) \) and \( \frac{1}{x} \), respectively, and we let

\[
\sigma'_n = \sigma_n \left(1 + \frac{2}{(2w-1)^2} \cdot \frac{\varepsilon}{(1-x)^2}\right), \quad \text{and} \quad \check{\delta}_n = \frac{\tilde{\delta}_n}{2w-1-\alpha^\varepsilon(w) + \sigma'_n} + \frac{\sigma'_n}{(2w-1-\alpha^\varepsilon(w))^2}.
\]

Hence, we observe \( \lim_{n \to \infty} \check{\delta}_n = 0 \) and conclude the lower bound:

\[
Q(w) \geq \frac{1}{2w-1-\alpha^\varepsilon(w)} - \check{\delta}_n.
\]

Finally, we let \( \delta_n = \max\{\check{\delta}_n, \sigma'_n\} \). Then, combining Eqs. (16) and (21), we get

\[
\frac{1}{2w-1-\alpha^\varepsilon(w)} - \delta_n \leq Q(w) \leq \frac{1}{2w-1-\alpha^\varepsilon(w) - \bar{\beta}(w)} + \delta_n.
\]

This completes the proof of Theorem 1.

Observe that Eq. (22) yields an asymptotically tight characterization of the minimax prediction risk in the context of complete graphs. Specifically, letting \( n \to \infty \), we have

\[
\frac{1}{2w-1-\alpha^\varepsilon(w)} \leq Q(w) \leq \frac{1}{2w-1-\alpha^\varepsilon(w) - \bar{\beta}(w)}.
\]

### 7.3 Proof of Theorem 2: Water-Filling Strategy for Private Network Design

Previously, we have seen that on complete graphs, the agent can carefully design a strategy to harvest the intrinsic uncertainty and ensure an additive overhead. In the next two sections (Sections 7.3 and 8), we will build on these insights and extend the analysis and policy design to non-complete graphs. Subsequently, we will show that the additive overhead due to uncertainty can still be attained despite more complex topologies. In this section, we consider the problem of private network design, where the agent may construct the graph subject to an average degree constraint, and prove Theorem 2 in the process. In
constructed as follows. Moreover, there exists one edge from any vertex to any other clique so that all connected component functions as a complete graph, and the time to travel between different components is small. Next, we design the Clique Water-Filling Strategy to be applied on $\mathcal{G}(n, \bar{p}_n)$ and quantify its delay as a function of the delay of $\psi_{\bar{q}}^{wf}$ on a complete graph. To do so, we use a coupling argument in which we “freeze” time whenever the agent leaves the component containing her goal. Finally, we derive an upper bound on the prediction risk using a reduction argument to the complete graph setting. Specifically, we observe that if the agent announces which component contains her goal, then the adversary can ignore the rest of the graph and focus only on the announced component. Since this component is itself a complete graph, the maximum prediction risk under $\psi_{\bar{q}}^{wf}$ applies as an upper bound. We note that the graphs and the strategy developed in this subsection generalize the complete graphs and the Water-Filling Strategy, respectively.

We start with the following definition that constructs the graph $G$ with the desired properties.

**Definition 3** ($k$-Clique Graph). Fix $n, k \in \mathbb{N}$, and denote $m = \frac{n}{k}$. Let $G = (\mathcal{V}, \mathcal{E})$ be a $k$-clique graph with $n$ vertices. Then, $G$ is the union of $k$ complete graphs, each with $m$ vertices. Moreover, there exists one edge from any vertex to any other clique so that all of the $k$ cliques are connected to each other through all of their vertices. Specifically, $G$ is constructed as follows.

1. For each $i \in \{1, \ldots, k\}$, construct a complete graph $G_i = (\mathcal{V}_i, \mathcal{E}_i)$ where $\mathcal{V}_i = \{v^i_1, \ldots, m^i\}$.
2. Define $\mathcal{E} = \{(v^i_j, v^j_i) : v \in \{1, \ldots, m\} \text{ and } 1 \leq i < j \leq k\}$.
3. Let $\mathcal{V} = \bigcup_{i=1}^{k} \mathcal{V}_i$ and $\mathcal{E} = \mathcal{E} \cup (\bigcup_{i=1}^{k} \mathcal{E}_i)$.
4. Set $G = (\mathcal{V}, \mathcal{E})$ so that $|\mathcal{V}| = n$ and $|\mathcal{E}| = \frac{n}{2} \left(\frac{n}{k} + k - 2\right)$.

If the average degree of the network is constrained to be at most $\bar{p}_n$, then one can choose the number of cliques, $k$, so as to have $\frac{n}{k} + k - 2 \leq \bar{p}_n$. Therefore, we let $\mathcal{G}(n, \bar{p}_n)$ be the family of $k$-clique graphs where $k$ satisfies $\frac{n}{k} + k - 2 \leq \bar{p}_n$. Note that a complete graph is a 1-clique graph.

Next, we exhibit how to design the Clique Water-Filling Strategy, which can be implemented on a $k$-clique graph $G$, and denote this strategy by $\psi^k$. The main idea is to apply the Water-Filling Strategy whenever the state is inside the same clique with the goal and to “freeze” time whenever the state is outside. Observe that the original Water-Filling Strategy is obtained as a special case of the Clique Water-Filling Strategy on a 1-clique graph.

**Definition 4** (Clique Water-Filling Strategy). Fix a target risk level $\bar{q} \in [0, 1]$. Let $G = (\mathcal{V}, \mathcal{E})$ be a $k$-clique graph with $n$ vertices and let $\mathcal{V}_D$ denote the set of vertices in the clique containing the goal. The strategy, $\psi^k_{\bar{q}}$, on $G$ is defined as follows.

1. **Counter.** Initialize a counter $C_0^{wf} = 0$. At each time $t \geq 1$, increment $C_t^{wf} = C_{t-1}^{wf} + 1$ if $X_t \in \mathcal{V}_D$.
(2) Actions. If the current state is outside $V_D$, travel to $V_D$. If the current state is in $V_D$, execute the next action under the Water-Filling Strategy, $\psi_{\tilde{V}}^{wf}$, where time is indexed by $C_{\tilde{V}}^{wf}$.

(a) If $X_t \in V_D$, set $a_t = \tilde{a}_{C_t^{wf}}$ where $\tilde{a}_{C_t^{wf}}$ is the meta action at time $C_t^{wf}$ under $\psi_{\tilde{V}}^{wf}$.

(b) If $X_t = v^t$ with $V_t \neq V_D$, set $a_t = v^{V_D}$ to go back to $V_D$.

Now, we quantify the delay under $\psi^k$ in terms of the expected intentional goal-hitting time under the Water-Filling Strategy on a complete graph with $n$ vertices. The proof is given in Appendix E.8. Intuitively, the delay under $\psi^k$ is the sum of two terms: the expected time spent in the clique containing the goal and the expected time spent outside. To prove Proposition 6, we begin by providing an upper bound on the expected total time the agent spends outside $V_D$. Then, we observe that the time spent inside $V_D$ is less than the expected intentional goal-hitting time under $\psi_{\tilde{V}}^{wf}$, where the time is indexed by the counter $C_{\tilde{V}}^{wf}$ and hence, conclude the proof.

Proposition 6. Fix horizon $K \in \mathbb{N}$ and target risk level $\bar{q} \in [0, 1]$. Fix a degree sequence $\{\bar{\rho}_n\}_{n \in \mathbb{N}}$ such that $\bar{\rho}_n \leq n$ and $\bar{\rho}_n \gg \sqrt{n}$. Fix graph size $n \in \mathbb{N}$ and clique number $k \in \mathbb{N}$ such that $\frac{n}{k} + k - 2 \leq \bar{\rho}_n$. Let $G = (V, E)$ be a $k$-clique graph with $|V| = n$. Under the Clique Water-Filling Strategy, $\psi_{\tilde{V}}^k$, the expected delay satisfies

$$\mathbb{E}(T^k) \leq \mathbb{E}(T_{IH}^{wf}) \left(1 + \frac{(k-1)k}{(1-\varepsilon)(n + (k-2)k + \varepsilon k)}\right),$$

where $T_{IH}^{wf}$ is the intentional goal-hitting time under the Water-Filling Strategy $\psi_{\tilde{V}}^{wf}$, on a complete graph with $\frac{n}{k}$ vertices.

The following result illustrates that the maximal prediction risk under $\psi_{\tilde{V}}^k$ is indeed at most $\bar{q}$. The proof relies on the observation that if the agent announces the clique containing her goal, then the game will evolve as if it is on a complete graph with $\frac{n}{k}$ vertices, as such the maximal prediction risk for the relevant Water-Filling Strategy will apply. The proof is given in Appendix E.9.

Proposition 7. Fix a sequence of horizons $\{K_n\}_{n \in \mathbb{N}}$ such that $1 \ll K_n \ll \sqrt{n}$. Fix a degree sequence $\{\bar{\rho}_n\}_{n \in \mathbb{N}}$ such that $\bar{\rho}_n \leq n$ and $\bar{\rho}_n \gg \sqrt{n}$. Fix graph size $n \in \mathbb{N}$ and clique number $k \in \mathbb{N}$ such that $\frac{n}{k} + k - 2 \leq \bar{\rho}_n$. Let $G = (V, E)$ be a $k$-clique graph with $|V| = n$. Under the Clique Water-Filling Strategy, $\psi_{\tilde{V}}^k$, the maximal prediction risk satisfies

$$q^*(\psi_{\tilde{V}}^k) \leq \bar{q} + \delta_n,$$

where $\delta_n = \frac{\bar{w}}{K_n} + 1 - \left(1 - \frac{k}{n}\right)^{K_n}$, and $\lim_{n \to \infty} \delta_n = 0$.

We are now ready to prove Theorem 2 by combining the results proven so far.

Proof of Theorem 2. Let $G = \mathcal{G}(n, \bar{\rho}_n)$ be a $k$-clique graph with $n$ vertices, for some $k \in \mathbb{N}$ such that $\frac{n}{k} + k - 2 \leq \bar{\rho}_n$. Then, consider the agent strategy $\psi_{\tilde{V}}^k$, that is, the Clique Water-Filling Strategy with target risk level $\bar{q} \in [0, 1]$. By Proposition 6, the expected delay under strategy $\psi_{\tilde{V}}^k$ satisfies

$$\mathbb{E}(T^k) \leq \mathbb{E}(T_{IH}^{wf}) \left(1 + \frac{(k-1)k}{(1-\varepsilon)(n + k^2 - 2k + \varepsilon k)}\right) = \mathbb{E}(T_{IH}^{wf})(1 + x).$$
Then, to obtain $E(T^k) \leq w$, it suffices to have $E(T^w_{\text{IH}})^{(1+x)} \leq w$. By Proposition 3, for $E(T^w_{\text{IH}}) \leq y$, it suffices to set $\tilde{q} = \frac{1}{2y - 1 - c^\epsilon}$ where $c^\epsilon = \frac{\epsilon}{(1-\epsilon)^2}$. Thus, letting $y = \frac{w}{1+x}$, Proposition 7 implies

$$Q(w) \leq q^*(\psi^k) \leq \tilde{q} + \delta_n = \frac{1}{2w\rho_n^c(k) - 1 - c^\epsilon} + \delta_n,$$

where $\rho_n^c(k) = \frac{1}{1+x} = \frac{(1-\epsilon)(n+k^2-2k)+\epsilon k}{n+2k^2-3k-\epsilon(n+k^2-3k)}$. Finally, by the assumption $\tilde{p}_n \gg \sqrt{n}$, for each $k \in \mathbb{N}$, we have $\lim_{n \to \infty} \rho_n^c(k) = 1$. This completes the proof of Theorem 2. 

Note that as $k$ increases, the average degree of the network with design parameter $k$ decreases while the prediction risk grows. Therefore, by tuning the value of $k$, one can design a private network with the most efficient cost and privacy tradeoff within this family. Finally, letting $n \to \infty$, we obtain the following upper bound on the minimax prediction risk

$$Q(w) \leq \frac{1}{2w - 1 - c^\epsilon},$$

implying that the performance on complete graphs can be asymptotically achieved by a careful design of the network topology.

8 Random Walk Water-Filling Strategy on Non-Complete Graphs

We have demonstrated in the preceding section that when the agent has the ability to design her network, she can achieve a (near) optimal prediction risk vs. delay tradeoff by carefully harvesting the intrinsic uncertainty, even in a relatively sparse graph. In certain real-world applications, however, the agent may not be able to design the network. In this section, we examine a broader class of non-complete graphs and develop the Random Walk Water-Filling Strategy, under which the prediction risk only depends on the maximum and minimum degrees of a given graph. Then, in order to ensure desirable minimax prediction risk, we construct a sequence of families of graphs, $\{G_n(p)\}_{n \in \mathbb{N}}$, on which the Random Walk Water-Filling Strategy admits both low prediction risk and delay. Consequently, we prove Theorem 3 by showing that for graphs in $\{G_n(p)\}_{n \in \mathbb{N}}$ an additive delay overhead can still be attained. Moreover, we prove that an Erdős-Rényi random graph with $n$ vertices and edge density $p \in [0, 1]$ belongs to the family $G_n(p)$ with high probability as $n \to \infty$. In a way, this suggests that “typical” graphs with average degree $pn$ under the Erdős-Rényi model still enjoy low minimax prediction risk as was the case in complete graphs. Finally, in Appendix G, we prove Theorem 5 which extends the analysis to sparse graphs where the edge density $p_n$ may approach zero as $n \to \infty$.

Compared to the topologies in private network design (Section 7.3), the lack of strong symmetries in the graphs in $\{G_n(p)\}_{n \in \mathbb{N}}$ makes it significantly more difficult to design a private agent strategy. In particular, the neighborhood of the goal vertex must also be obfuscated, for it can be highly revealing of the goal itself. This was not the case in a complete graph, or the $k$-clique graph, where a large number of vertices share the same neighborhood as the goal vertex. To mitigate this issue, we will combine the original Water-Filling Strategy with a random walk, so that the adversary cannot predict in what neighborhoods an attempt to reach the goal has been made. As a result of such obfuscation of the neighborhoods, we will show that, surprisingly, the use of a random walk leads to little additional overhead.
8.1 Random Walk Water-Filling Strategy

We begin by defining the Random Walk Water-Filling Strategy, $\psi^p$, which can be used on any undirected graph. In a nutshell, the agent performs a random walk on the graph, and attempts to implement a step in the original Water-Filling Strategy whenever the random walk takes her to a neighboring vertex of the goal. Following the formal definition, we characterize the maximal prediction risk of $\psi^p$ on any undirected graph, in terms of its minimum and maximum degrees. In the subsequent section, we use this characterization to derive stronger performance guarantees for the Random Walk Water-Filling Strategy on the specific families of graphs, $\{G_n(p)\}_{n \in \mathbb{N}}$, and show that it achieves additive overhead.

**Definition 5 (Random Walk Water-Filling Strategy).** Fix a target risk level $\bar{q} \in [0, 1]$. Let $G = (\mathcal{V}, \mathcal{E})$ be a graph with $n$ vertices. The strategy, $\psi^p_{\bar{q}}$, on $G$ is defined as follows.

1. **Counter.** Initialize a counter $C_{wf}^t = 0$. At each $t \geq 1$, increment $C_{wf}^t = C_{wf}^{t-1} + 1$ if $X_t \in \mathcal{E}(D)$, i.e., if $X_t$ is in the neighborhood of the goal $D$.

2. **Actions.** If the current state is outside $\mathcal{E}(D)$, perform a random walk by selecting the next action uniformly at random from the set of neighboring vertices. If the current state is a neighbor of $D$, i.e., $X_t \in \mathcal{E}(D)$, execute the next action under the Water-Filling Strategy, $\psi^w_{\bar{q}}$, where time is indexed by the counter $C_{wf}^t$. In summary,

   (a) If $X_t \in \mathcal{E}(D)$, set $a_t = \bar{a}_{C_{wf}^t}$ where $\bar{a}_{C_{wf}^t}$ is the meta action at time $C_{wf}^t$ under $\psi^w_{\bar{q}}$.

   (b) If $X_t \not\in \mathcal{E}(D)$, set $a_t = v$ where $v \sim \text{Unif}(\mathcal{E}(X_t))$.

**Proposition 8.** Fix graph size $n \in \mathbb{N}$, horizon $K \in \mathbb{N}$ and target risk level $\bar{q} \in [0, 1]$. Let $G = (\mathcal{V}, \mathcal{E})$ be a graph with $|\mathcal{V}| = n$. Under the Random Walk Water-Filling Strategy, $\psi^p_{\bar{q}}$, the maximal prediction risk satisfies

$$q^*(\psi^p_{\bar{q}}) \leq \bar{q} \frac{\Sigma_G}{n} + 1 - \left(1 - \frac{1}{\Delta_G}\right)^K + f_K,$$

where $f_K = \frac{\bar{u}}{K}$, $\lim_{K \to \infty} f_K = 0$, and $\Delta_G$ and $\Sigma_G$ are the minimum and maximum degrees of $G$, i.e., $\Delta_G = \min_{v \in \mathcal{V}} |\mathcal{E}(v)|$ and $\Sigma_G = \max_{v \in \mathcal{V}} |\mathcal{E}(v)|$.

8.2 Graph Family $G_n(p)$

In the previous section, we have seen that the prediction risk of $\psi^p$ on a non-complete graph depends on its minimum and maximum degrees. We now formally define a parameterized family of graphs $\mathcal{F}_n(p, \gamma)$ with two key properties, which ensure that the minimax prediction risk is low and the delay is small.
Definition 6 (Family \(F_n(p, \gamma)\) of Graphs). Fix \(p \in [0,1]\) and \(\gamma \in [0,1]\). For each \(n \in \mathbb{N}\), we define \(F_n(p, \gamma)\) as the set of all graphs \(G = (V, E)\) with \(n\) vertices that satisfy the following properties.

1. **Regularity.** The degrees of all vertices are upper- and lower-bounded by \(pn(1 + \gamma)\) and \(pn(1 - \gamma)\), respectively, i.e., \(pn(1 - \gamma) \leq |E(v)| \leq pn(1 + \gamma)\) for all \(v \in V\).

2. **Neighborhood Overlap.** For any pair \(u, v \in V\), at least \(p(1 - \gamma)\) fraction of the neighborhood of \(v\) intersects the neighborhood of \(u\), i.e., \(|E(u) \cap E(v)| / |E(v)| \geq p(1 - \gamma)\).

Intuitively, the Neighborhood Overlap property guarantees that the delay is small since a random walk will frequently visit the neighborhood of the goal vertex. Similarly, the Regularity property implies that the degrees of vertices are concentrated around \(pn\) which in light of Proposition 8 will allow us to show that the prediction risk is low.

Next, we will set the value of \(\gamma\) in \(F_n(p, \gamma)\) to define our main object of interest, the sequence of families \(\{G_n(p)\}_{n \in \mathbb{N}}\).

Definition 7 (Family \(G_n(p)\) of Graphs). Fix \(p \in [0,1]\). Let \(\{\gamma_n\}_{n \in \mathbb{N}}\) be a sequence such that \(\gamma_n \in [0, \frac{1}{2}]\); \(\gamma_n \geq \sqrt{\log n/n}\) and \(\lim_{n \to \infty} \gamma_n = 0\). For each \(n \in \mathbb{N}\), we define \(G_n(p) = F_n(p, \gamma_n)\).

Note that in the remainder of Section 8, we will restrict our focus on the family \(G_n(p)\) where \(\gamma_n\) is defined as in Definition 7. Recall that under the Erdős-Rényi random graph model, a graph with \(n\) vertices is constructed by independently connecting each pair of vertices by an edge with probability \(p\). Lemma 3 shows that with high probability, a graph generated using the Erdős-Rényi random graph model with edge probability \(p\) belongs to the family \(G_n(p)\). The proof relies on an application of concentration inequalities and is given in Appendix E.11. We note that Lemma 3 proves Theorem 3(a).

Lemma 3. Fix edge density \(p \in [0,1]\). Fix graph size \(n \in \mathbb{N}\) and consider the family of graphs \(G_n(p)\). Let \(G\) be an Erdős-Rényi random graph with \(n\) vertices and edge probability \(p\). Then, there exists a sequence \(\{\theta_n\}_{n \in \mathbb{N}}\), with \(\lim_{n \to \infty} \theta_n = 0\), such that for every \(n \in \mathbb{N}\), \(\mathbb{P}(G \in G_n(p)) \geq 1 - \theta_n\) holds.

8.3 Proof of Theorem 3(b)

We now turn to the proof of Theorem 3(b). The proof will leverage two key properties that define the graphs in \(\{G_n(p)\}_{n \in \mathbb{N}}\). First, we use the Regularity property to derive an upper bound on the prediction risk, and show that it can be bounded by the risk under the Water-Filling Strategy scaled by a factor of \(p\). Next, we use the Neighborhood Overlap property to show that an agent using a random walk will “stumble” upon the neighborhood of her goal every \(\frac{1}{p}\) time steps on average. We subsequently combine this observation with a “time-stretching” argument to establish that the delay under the agent strategy \(\psi^p\) is upper-bounded by that of \(\psi^{\text{wf}}\) on a complete graph, but slowed down by a factor of \(\frac{1}{p}\). Combining the characterizations of risk and delay completes the proof. (Note that the two scaling factors \(\frac{1}{p}\) and \(p\) neutralize one another.)

We begin by analyzing the prediction risk under \(\psi^p\) and show that the maximal prediction risk becomes asymptotically upper-bounded by \(pq\) as the number of vertices, \(n\), grows. The proof appears in Appendix E.12.
Proposition 9. Fix a sequence of horizons \(\{K_n\}_{n \in \mathbb{N}}\) such that \(1 \ll K_n \ll n\). Fix graph size \(n \in \mathbb{N}\), edge density \(p \in [0, 1]\) and target risk level \(\bar{q} \in [0, 1]\). Let \(G = (\mathcal{V}, \mathcal{E})\) be a graph from the family \(\mathcal{G}_n(p)\). Under the Random Walk Water-Filling Strategy, \(\psi^p_{\bar{q}}\), the maximal prediction risk satisfies

\[
q^*(\psi^p_{\bar{q}}) \leq \bar{q}p(1 + \gamma_n) + \delta'_n,
\]

where \(\delta'_n = \frac{w}{K_n} + 1 - \left(1 - \frac{1}{pn(1 - \gamma_n)}\right)^{K_n}\), and \(\lim_{n \to \infty} \delta'_n = 0\).

We next turn to the delay. The following result establishes that the delay under \(\psi^p\) is upper-bounded by the expected intentional goal-hitting time under the Water-Filling Strategy on a complete graph, scaled by \(\frac{1}{p}\). Intuitively, the vertices of the graph \(G\) are on average connected to \(p\) fraction of the vertices so that under \(\psi^p\), the agent has an opportunity to attempt her goal \(p\) fraction of the time. Using a coupling argument we show that \(\mathbb{E}(T^p_{IH})\) slowed down by \(\frac{1}{p}\) gives an upper bound for \(\mathbb{E}(T^p_{IH})\). The proof is in Appendix E.13.

Proposition 10. Fix graph size \(n \in \mathbb{N}\), horizon \(K \in \mathbb{N}\), edge density \(p \in [0, 1]\) and target risk level \(\bar{q} \in [0, 1]\). Let \(G = (\mathcal{V}, \mathcal{E})\) be a graph from the family \(\mathcal{G}_n(p)\). Under the Random Walk Water-Filling Strategy, \(\psi^p_{\bar{q}}\), the expected delay satisfies

\[
\mathbb{E}(T^p) \leq \frac{\mathbb{E}(T^p_{IH}))}{p(1 - \gamma_n)} + 1,
\]

where \(T^p_{IH})\) is the intentional goal-hitting time under the Water-Filling Strategy, \(\psi^p_{\bar{q}}\), on a complete graph with \(n\) vertices.

Recall that Lemma 3 proves Theorem 3(a). To prove Theorem 3(b), we combine Propositions 9 and 10. The proof is provided in Appendix E.14 and establishes that

\[
\mathcal{Q}(w) \leq \frac{1}{w - \lambda\varepsilon(p)} + \delta_n,
\]

where \(\lambda\varepsilon(p) = 1 + \frac{1}{p} + \frac{\varepsilon}{(1 - \varepsilon)p}\), and

\[
\delta_n = \frac{\gamma_n}{\sqrt{\frac{\varepsilon}{(1 - \varepsilon)p}} - \frac{\varepsilon}{(1 - \varepsilon)p}} + 1 - \left(1 - \frac{1}{pn(1 - \gamma_n)}\right)^{K_n} + \frac{w}{K_n}.
\]

Note that as we let \(n\) and \(K\) grow, the upper bound becomes

\[
\mathcal{Q}(w) \leq \frac{1}{w - 1 - \frac{1}{p} - \frac{\varepsilon}{(1 - \varepsilon)p}},
\]

showing that an additive uncertainty overhead is still achievable, even when the graph is generated from a random graph model. In Appendix G, we generalize this result to sparse non-complete graphs and compare the performance with that under network design. In particular, we show that even though additive overhead can still be attained, Theorem 2 provides a stronger performance guarantee as it allows for a more sparse network and a smaller delay budget.
9 Numerical Experiments

While most of our main results concern the privacy vs. efficiency tradeoff in large graphs, in this section we present numerical studies and simulations on finite graphs with moderate size. We examine synthetic as well as real-world networks with moderate size to study the tradeoff.

We first simulate the Water-Filling Strategy on a complete graph with 200 vertices. The results are given in Figure 2, where the agent uses the Water-Filling Strategy with $\bar{q} = \frac{1}{2w-1-\epsilon^2}$, where $c^2 = \frac{\epsilon}{(1-\epsilon^2)^2}$, for a range of delay budget values, $w$; each subplot corresponds to a different value of noise level, $\epsilon$. Each marker on the solid curve with diamonds corresponds to a fixed value of the delay budget and the prediction risk is computed by averaging over 500 goal node realizations with 1000 trajectory realizations each. The adversary strategy is assumed to consist of making a prediction at a randomly sampled time over the first $t^*$ periods, where $t^*$ is defined in Definition 2. The dash-dot line plots the lower bound in Theorem 4, i.e., $\frac{1}{2w+1}$.

![Figure 2](image_url)

Figure 2: An illustration of the prediction risk as a function of the delay under the Water-Filling Strategy on a complete graph with 200 vertices.

Note that as $\epsilon$ increases, the prediction risk also increases for any delay budget since the agent is required to choose higher $\bar{q}$ values to ensure similar delay values. While our main result consisted of an asymptotically tight characterization of the minimax prediction risk, these results illustrate that our bounds are relatively tight even for a finite graph with
We now turn to general graphs and conduct numerical experiments on a real-world peer-to-peer network. We use the snapshot of the Gnutella peer-to-peer file sharing network from August 4, 2002 (Ripeanu and Foster (2002)). The Gnutella network was one of the first examples of peer-to-peer networks and had a substantial user base while it was operational (Ripeanu and Foster (2002)). The network consists of 10876 vertices corresponding to the hosts in the Gnutella network and 39994 edges representing the connections between them. Furthermore, the network has diameter $d_G = 10$.

Figure 3 illustrates the simulation results on the Gnutella network where the agent uses the segment-based strategy with greedy shortest path routine (Section 6), for varying values of the delay budget $w$. Each subplot depicts the simulation results under a different noise level, $\varepsilon$. Each marker corresponds to a fixed delay budget value and the prediction risk is computed by averaging over $S = 1000$ realizations of the trajectory for $M = 500$ goal vertex realizations. The dash-dot line plots the lower bound in Theorem 4, i.e., $\frac{1}{2w+\varepsilon}$. Further, the solid curve with diamond markers demonstrates the scenario where the adversary makes a prediction at a randomly sampled time over the length of the trajectory. A step-by-step description of the simulation procedure for general graphs is given in Algorithm 1 in Appendix C.

Figure 3: An illustration of the prediction risk as a function of the delay budget under the stochastic segment-based strategy on the Gnutella network.

First, we note that the prediction risk is inversely proportional to the delay budget in both subplots. Further, as the noise level $\varepsilon$ increases, the agent has to choose larger delay budget values. Nevertheless, the prediction risk slightly decreases with increasing noise levels for a fixed delay budget. This is because the adversary strategy chooses a time over the whole trajectory and as the noise level increases, a greater proportion of the vertices within the trajectory are chosen by nature. Compared to the simulations on the complete graph, the bounds are less tight for general graphs. The numerical experiments suggest that for moderate sized graphs, the tradeoff retains the same qualitative characteristics described by our theoretical results. However, as expected, the bounds have widened in the numerical studies compared to the theoretical results.
10 Conclusion

This paper proposes a framework, \( \varepsilon \)-Noisy Goal Prediction Game, to analyze privacy-preserving decision policies in a dynamic decision-making problem under uncertainty and to study the optimal tradeoff between delay and privacy on different network topologies. Our main result establishes a tight characterization of the minimax prediction risk, in the context of complete graphs, and provides an optimal agent strategy achieving the optimal tradeoff between delay and privacy. In particular, we show that the agent has to incur a strictly positive but additive overhead due to uncertainty. We further consider the problem of private network design and construct a family of networks attaining the most efficient delay and privacy tradeoff subject to an average degree constraint. Next, we analyze graphs generated using the Erdős-Rényi random graph model and prove that similar performance can be achieved. Finally, we provide a strategy allowing the agent to operate on any graph under any level of intrinsic uncertainty, and illustrate that natural extensions of policies designed under the perfect-control assumptions perform poorly under uncertainty.

There is a number of interesting directions for future research. The current paper assumes that the adversary is only able to observe the effects of the agent’s actions. One may also want to extend our analysis to different models of the adversary. For instance, one could model an adversary who observes the actions taken by the agent but not their effects. This scenario would be a more realistic model for applications such as medical treatments, where often the effect of the chosen treatment is unobserved. Alternatively, one could allow an adversary who is able to observe both the action and its effect. It would be interesting to see how the overhead on privacy is affected depending on the specific assumptions on the structure of the information that the adversary possesses.

References

Allioux, R. (2018). How hard is it for an ai to detect ships on satellite images?
Augenblick, N. and Bodoh-Creed, A. (2018). To reveal or not to reveal: Privacy preferences and economic frictions. Games and Economic Behavior, 110:318–329.
Bertsekas, D. P. and Tsitsiklis, J. N. (1991). An analysis of stochastic shortest path problems. Mathematics of Operations Research, 16(3):580–595.
Blum, A., Morgenstern, J., Sharma, A., and Smith, A. (2015). Privacy-preserving public information for sequential games. In Proceedings of the 2015 Conference on Innovations in Theoretical Computer Science, pages 173–180. ACM.
Calmon, F. P., Makhdoumi, A., and Médard, M. (2015). Fundamental limits of perfect privacy. In 2015 IEEE International Symposium on Information Theory (ISIT), pages 1796–1800. IEEE.
Chaudhuri, K., Monteleoni, C., and Sarwate, A. D. (2011). Differentially private empirical risk minimization. Journal of Machine Learning Research, 12(Mar):1069–1109.
Corrigan-Gibbs, H., Boneh, D., and Mazières, D. (2015). Riposte: An anonymous messaging system handling millions of users. In 2015 IEEE Symposium on Security and Privacy, pages 321–338. IEEE.
Croucher, J. S. (1978). A note on the stochastic shortest-route problem. Naval Research Logistics Quarterly, 25(4):729–732.
Cummings, R., Echenique, F., and Wierman, A. (2016). The empirical implications of privacy-aware choice. Operations Research, 64(1):67–78.
De Montjoye, Y.-A., Hidalgo, C. A., Verleysen, M., and Blondel, V. D. (2013). Unique in the crowd: The privacy bounds of human mobility. Scientific reports, 3:1376.
Dingledine, R., Mathewson, N., and Syverson, P. (2004). Tor: The second-generation onion router. Technical report, Naval Research Lab Washington DC.

Dwork, C., Roth, A., et al. (2014). The algorithmic foundations of differential privacy. Foundations and Trends® in Theoretical Computer Science, 9(3–4):211–407.

Eaton, J. and Zadeh, L. (1962). Optimal pursuit strategies in discrete-state probabilistic systems. Journal of Basic Engineering, 84(1):23–29.

Fanti, G., Kairouz, P., Oh, S., and Viswanath, P. (2015). Spy vs. spy: Rumor source obfuscation. In ACM SIGMETRICS Performance Evaluation Review, volume 43, pages 271–284. ACM.

Flajolet, A., Blandin, S., and Jaillet, P. (2017). Robust adaptive routing under uncertainty. Operations Research, 66(1):210–229.

Goel, S., Robson, M., Polte, M., and Sirer, E. (2003). Herbivore: A scalable and efficient protocol for anonymous communication. Technical report, Cornell University.

Gradwohl, R. and Smorodinsky, R. (2017). Perception games and privacy. Games and Economic Behavior, 104:293–308.

Gupta, A., Roth, A., and Ullman, J. (2012). Iterative constructions and private data release. In Theory of cryptography conference, pages 339–356. Springer.

Issa, I., Kamath, S., and Wagner, A. B. (2016). An operational measure of information leakage. In 2016 Annual Conference on Information Science and Systems (CISS), pages 234–239. IEEE.

Jaillet, P., Qi, J., and Sim, M. (2016). Routing optimization under uncertainty. Operations research, 64(1):186–200.

Jain, P., Kothari, P., and Thakurta, A. (2012). Differentially private online learning. In Conference on Learning Theory, pages 24–1.

Kim, S., Lewis, M. E., and White, C. C. (2005). Optimal vehicle routing with real-time traffic information. IEEE Transactions on Intelligent Transportation Systems, 6(2):178–188.

Kurose, J. F. and Ross, K. W. (2013). Computer networking: a top-down approach. Addison Wesley.

Lindell, Y. and Pinkas, B. (2009). Secure multiparty computation for privacy-preserving data mining. Journal of Privacy and Confidentiality, 1(1).

Mayer, J., Mutchler, P., and Mitchell, J. C. (2016). Evaluating the privacy properties of telephone metadata. Proceedings of the National Academy of Sciences, 113(20):5536–5541.

Mestre, A. M., Oliveira, M. D., and Barbosa-Póvoa, A. P. (2015). Location-allocation approaches for hospital network planning under uncertainty. European Journal of Operational Research, 240(3):791–806.

Miller-Hooks, E. (2001). Adaptive least-expected time paths in stochastic, time-varying transportation and data networks. Networks: An International Journal, 37(1):35–52.

Nagurney, A., Shukla, S., Nagurney, L. S., and Saberi, S. (2018). A game theory model for freight service provision security investments for high-value cargo. Economics of transportation, 16:21–28.

Perkins, C. E. et al. (2001). Ad hoc networking, volume I. Addison-wesley Reading.

Radaelli, L., Sapiezynski, P., Houssiau, F., Shmueli, E., and de Montjoye, Y.-A. (2018). Quantifying surveillance in the networked age: Node-based intrusions and group privacy. arXiv preprint arXiv:1803.09007.

Ratnam, G. (2019). Your car is watching you. who owns the data? Roll Call.

Ripeanu, M. and Foster, I. (2002). Mapping the gnutella network: Macroscopic properties of large-scale peer-to-peer systems. In international workshop on peer-to-peer systems, pages 85–93. Springer.

Rodriguez, D. (2018). Satellite-based ground truth for parking availability.
Sezer, V., Bandyopadhyay, T., Rus, D., Frazzoli, E., and Hsu, D. (2015). Towards autonomous navigation of unsignalized intersections under uncertainty of human driver intent. In *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*, pages 3578–3585. IEEE.

Shokri, R. and Shmatikov, V. (2015). Privacy-preserving deep learning. In *Proceedings of the 22nd ACM SIGSAC conference on computer and communications security*, pages 1310–1321. ACM.

The Economist (2019). The rise of the financial machines.

Tossou, A. C. and Dimitrakakis, C. (2016). Algorithms for differentially private multi-armed bandits. In *AAAI*, pages 2087–2093.

Tsitsiklis, J. N. and Xu, K. (2018). Delay-predictability trade-offs in reaching a secret goal. *Operations Research*, 66(2):587–596.

Tsitsiklis, J. N., Xu, K., and Xu, Z. (2018). Private sequential learning. *arXiv preprint arXiv:1805.02136*.

Valentino-DeVries, J., Singer, N., Keller, M. H., and Krolik, A. (2018). Your apps know where you were last night, and they’re not keeping it secret. *NY Times*.

Vu, Q. H., Lupu, M., and Ooi, B. C. (2009). *Peer-to-peer computing: Principles and applications*. Springer Science & Business Media.

Wang, L., Yang, L., Gao, Z., Li, S., and Zhou, X. (2016). Evacuation planning for disaster responses: A stochastic programming framework. *Transportation research part C: emerging technologies*, 69:150–172.

Warren, J. (2012). Bitmessage: A peer-to-peer message authentication and delivery system. *white paper (27 November 2012), https://bitmessage.org/bitmessage.pdf*.

Xu, K. (2018). Query complexity of bayesian private learning. In *Advances in Neural Information Processing Systems*, pages 2432–2441.
A  Example Instance of the $\varepsilon$-Noisy Goal Prediction Game

Suppose that $G$ is as in Figure 4. The agent starts at vertex 1 and her secret goal is $D = 5$. The agent’s trajectory is $1 \rightarrow 2 \rightarrow 3 \rightarrow 5$ whereas her strategy is given by the sequence of actions $a_1 = 3$, $a_2 = 3$, $a_3 = 5$. The sequence of Bernoulli trials is given by $B_1 = B_3 = 1$ and $B_2 = 0$. Note that the state of the agent at time $t = 2$ being 2 even though her action was vertex 3 exemplifies a case in which the agent ends up at a different vertex due to noise. Assume that the adversary predicts at time $t = 2$ that the goal vertex is 3, i.e., $\hat{D}_2 = 3$. Since $\hat{D}_2 \neq D$, the agent wins the game with a goal-hitting time of $T = 4$. Nonetheless, if the adversary had made a prediction $\hat{D}_t = 5$ at any time $t \leq 4$, then he would have won.

![Figure 4: An example of the $\varepsilon$-Noisy Goal Prediction Game.](image)

B  Example Strategy that Reveals the Goal

To illustrate the importance of the independence between the trajectory and the intentional goal-hitting time, we describe here a simple example in which the agent’s actions are actually signaling what her goal is. Consider a complete graph in which the vertices are labeled with numbers $1, ..., n$ and that the goal vertex is vertex $i$. Suppose the agent’s strategy is to visit vertex $i - 1$ in period 1 and to do a random walk over the vertices starting at period 2. Then, the adversary can correctly guess the goal vertex, conditional on vertex $i - 1$ being successfully visited, and the prediction risk becomes $1 - \varepsilon$. The mistake in this strategy is that the action of the agent is revealing her goal even when she is not traveling towards it. This simple example demonstrates when the trajectory is not constructed independently from $T_{IH}$, the strategy itself might reveal the goal vertex and the adversary can indeed learn from the past actions of the agent.
C Simulation Procedure

Algorithm 1 Simulation Procedure for General Graphs

Fix $\varepsilon \in [0, 1]$. 

for $w \in [w_{\min}, w_{\max}]$ do
  Construct segment family, $S$.
  for $m \in \{1, \cdots, M\}$ do
    Sample a goal vertex, $D_m$.
    for $s \in \{1, \cdots, S\}$ do
      Play game $s$:
      (Agent) Sample path realization, $S(D_m)$.
      Traverse $S(D_m)$ using greedy shortest path routine.
      (Adversary) Sample time $t$ in trajectory.
      Predict $\hat{D}_s = X_t$.
      (Result) Determine the winner.
    end for
  end for
Compute average prediction risk for delay budget $w$.

D Potential Implementation with Bitmessage

In this section, we discuss how our policy for general graphs can be implemented in the context of anonymous messaging platforms. Specifically, we focus on the platform Bitmessage and explain how one of our proposed algorithms, the Stochastic Pre-Commit Strategy (see Section 6), can be implemented on this platform.

Bitmessage is a peer-to-peer communication protocol that enables users to send messages anonymously to each other (Warren (2012)). While a message may be sent to multiple users, each message is encrypted and can only be decoded by the private key of the actual recipient. The current Bitmessage protocol achieves recipient anonymity by broadcasting the message to all users. However, this approach would be difficult to scale as the network grows. First, sending the message to all nodes would cause high communication traffic overhead that grows with network size. Second, nodes in the network may be connected by a sparse and non-complete physical communication network. Hence, the overall performance of the messaging platform may be adversely affected due to the overhead in the network traffic.

To address these issues, our approach consists of a routing protocol that sends redundant messages only to a set of neighbors of the intended recipient instead of the entire network, where neighborhoods are defined with respect to some communication graph. However, the neighborhood is carefully sampled with a strong information-theoretic guarantee: an adversary who knows the graph and sees the neighborhood wouldn’t be able to infer which node therein is the true recipient.

The underlying Bitmessage network is a binary tree where each node itself is a cluster of multiple vertices. The nodes of the binary tree are called streams whereas each node within a stream represents a unique user. Whenever a message is being routed over the network,
it is first relayed to the target stream that is known to contain the address of the intended recipient. Next, the message is delivered to every user within the stream. Due to the structure of the network, each user can easily route messages across streams as long as they know the target stream number. In particular, the stream serves as a higher level decoy set that makes it difficult for nodes to be identifiable. Further, to achieve scalability the routers on the network only connect to some of the streams as opposed to all. Hence, proximity in the underlying network represents the ease of routing messages and the associated cost of network congestion.

Having explained the network structure and the routing protocol, we now describe how our policy can be implemented on Bitmessage. In particular, we consider the network of streams to be the graph \( G = (V, E) \) where the vertices are streams and the edges represent whether it is possible to directly relay a message from stream \( v \) to \( v' \in V \). Secondly, the decision maker chooses the delay budget \( w \) based on the level of efficiency that is desired to be achieved. Consequently, the decision maker can set \( r = w(1 - 2\varepsilon) - d_G \) where \( d_G \) is the diameter of the graph, \( \varepsilon \in [0, 1] \) represents the level of uncertainty and \( r \) denotes the length of each segment. We note that a segment of length \( r \) is a sequence of vertices \( s = (s_1, \ldots, s_r) \) such that \( (s_i, s_{i+1}) \in E \) for \( i = 1, \ldots, r - 1 \). Particularly, each vertex \( s_i \in s \) is a stream and \( r \) is the number of streams that are going to receive the message, including the fictitious recipients as well as the intended recipient.

Once the delay budget \( w \) is chosen, the decision maker needs to construct a family of segments, \( S \), using the underlying graph \( G \). The family \( S \) is a collection of sets of segments \( S_v \) for each stream \( v \in V \) and covers the graph in a uniform manner. More precisely, for each stream \( v \in V \), the set of segments containing \( v \) is denoted by \( S_v \). Further, the family \( S \) satisfies the following properties: (1) each segment \( s \in S \) has length \( r \), and (2) the size of \( S_v \) for each \( v \in V \) is at least \( r \). Then, for each message that is going to be routed over the network to user in stream \( D \), the protocol samples a segment that contains the target stream \( D \) and sends the message to each stream within the segment. Briefly, the strategy can be viewed as a random mapping \( S \) that takes as input the target stream \( D \), and outputs a set of streams that will receive the message. The mapping \( S \) is designed in such a way that: (a) the length of \( S(D) \) is at most \( w \), ensuring a small delay, and (b) \( S(D) \) does not significantly reveal \( D \), ensuring privacy.

Finally, there are two main sources of uncertainty in anonymous messaging platforms. First, there might be random link failures due to network congestion (cf. Perkins et al. (2001), Kurose and Ross (2013)). In the event of a link failure, the protocol needs to send the message again in a later time period in order to ensure its successful delivery. Second, in a peer-to-peer network some users might be offline at a given time period (cf. Warren (2012)). Consequently, the delivery of the message to an offline user might fail and the protocol might need to resend the message until it is delivered.

In order to account for these sources of uncertainty in the system, we propose the implementation of the Stochastic Pre-Commit Strategy jointly with the routing protocol. We note that the implementation of the Stochastic Pre-Commit Strategy corresponds to a routing protocol which tries sending a message to a given user repeatedly until it succeeds. Specifically, the chosen segment \( S(D) \) is treated as the sampled trajectory. Then, the message is routed to each stream within the segment using the Greedy Shortest Path (GSP) routine (See Section 6). We believe the implementation of our policies on anonymous messaging platforms is a promising direction of future work.
E Proofs

E.1 Proof of Proposition 1

Define the simple random walk \( \hat{Z}(t) = \sum_{i=1}^{t} Y_i \) with initial condition \( \hat{Z}(0) = 0 \). Suppose the random variables \( Y_i \) are independent and identically distributed with \( \mathbb{P}(Y_i = -1) = p \) and \( \mathbb{P}(Y_i = +1) = q \). Let \( \hat{T}_k = \min\{t \geq 0 : \hat{Z}(t) = -k\} \). We begin by proving the following lemma, which will be used shortly in the proof of Proposition 1.

Lemma 4. For any \( k \geq 1 \), we have \( \mathbb{E}(\hat{T}_k) = \frac{k}{2p-1} \).

Proof. We will show that the time between two consecutive stopping times, \( T_k \) and \( T_{k+1} \), has the same distribution with \( T_1 \). Then, we will use this result to express \( \mathbb{E}(T_k) \) in terms of \( \mathbb{E}(T_1) \).

Let \( k \geq 1 \). Consider the stopping time \( \hat{T}_k = \min\{t \geq 0 : \hat{Z}(t) = -k\} \). Rewriting \( \hat{T}_{k+1} = \hat{T}_{k+1} - \hat{T}_k + \hat{T}_k \), we have \( \mathbb{E}(\hat{T}_{k+1}) = \mathbb{E}(\hat{T}_{k+1} - \hat{T}_k) + \mathbb{E}(\hat{T}_k) \). Define \( D_k = \hat{T}_{k+1} - \hat{T}_k \). Now, we will show that for any \( k > 0 \), \( D_k \) and \( T_1 \) have the same distribution. First, observe that \( D_k \) can be written as:

\[
D_k = \hat{T}_{k+1} - \hat{T}_k = \min\{s \geq 0 : \hat{Z}(\hat{T}_k + s) - (k+1)\} = \min\{s \geq 0 : \hat{Z}(\hat{T}_k + s) - \hat{Z}(\hat{T}_k) = -1\},
\]

where the last equality follows from the definition of \( \hat{T}_k \). Since \( \hat{T}_k \) is a stopping time with respect to the natural filtration induced by \( \{Y_i\}_{i \in \mathbb{N}} \), we know that \( \{\hat{Z}(\hat{T}_k + t) - \hat{Z}(\hat{T}_k)\}_{t \in \mathbb{N}} \) has the same distribution with \( \{\hat{Z}(t)\}_{t \in \mathbb{N}} \).

Hence, using the fact that \( Y_i \) are independent and identically distributed, we obtain the following:

\[
\mathbb{P}(D_k = t) = \mathbb{P}(\min\{s \geq 0 : \hat{Z}(\hat{T}_k + s) - \hat{Z}(\hat{T}_k) = -1\} = t)
\]

\[
= \sum_{i=0}^{\infty} \mathbb{P}(\hat{T}_k = l) \mathbb{P}(D_k = t | \hat{T}_k = l)
\]

\[
= \sum_{i=0}^{\infty} \mathbb{P}(\hat{T}_k = l) \mathbb{P}\left( \min\left\{ s \geq 0 : \sum_{i=l}^{l+s} Y_i = -1 \right\} | \hat{T}_k = l \right)
\]

\[
= \sum_{i=0}^{\infty} \mathbb{P}(\hat{T}_k = l) \mathbb{P}(\hat{T}_1 = t | \hat{T}_k = l)
\]

\[
= \mathbb{P}(\hat{T}_1 = t).
\]

Therefore, \( \hat{T}_1 \overset{d}{=} D_k \) for any \( k \geq 1 \). This implies \( \mathbb{E}(\hat{T}_{k+1}) = \mathbb{E}(D_k) + \mathbb{E}(\hat{T}_k) = \mathbb{E}(\hat{T}_1) + \mathbb{E}(\hat{T}_k) \). Repeating the same argument, we obtain

\[
\mathbb{E}(\hat{T}_{k+1}) = \mathbb{E}(\hat{T}_1) + \mathbb{E}(\hat{T}_k) = \ldots = (k-1)\mathbb{E}(\hat{T}_1) + \mathbb{E}(\hat{T}_2) = (k+1)\mathbb{E}(\hat{T}_1).
\]

Finally, by conditioning on the first step, we get \( \mathbb{E}(\hat{T}_1) = 1 + 0p + q\mathbb{E}(\hat{T}_2) \). Then the equality \( \mathbb{E}(\hat{T}_2) = 2\mathbb{E}(\hat{T}_1) \) implies \( \mathbb{E}(\hat{T}_1) = 1 + 2q\mathbb{E}(\hat{T}_1) \). Thus, we conclude the proof with \( \mathbb{E}(\hat{T}_1) = \frac{1}{1-2q} = \frac{1}{2p-1} \) and \( \mathbb{E}(\hat{T}_k) = \frac{k}{2p-1} \). \( \Box \)

Recall that \( d(v, v') \) is the distance from vertex \( v \) to \( v' \) and the diameter of the graph is given by \( d_G = \max_{v, v' \in V} d(v, v') \). Let \( X(0) = v \) and define \( Z(t) = d(X(t), v') - d(v, v') \) for any \( t \geq 0 \). Also, denote the maximum degree of the graph by \( \Delta = \max_{v \in V} |\mathcal{E}(v)| \).
Also, recall that the GSP routine is the policy under which the agent always chooses the next vertex on the shortest path on $G$ from $X(t)$ to $v'$. Define $T = \min\{t \geq 0 : Z(t) = -d(v, v')\}$. Then, it follows that the length of the greedy shortest path between $v$ and $v'$ is the expected value of $T$ under GSP, i.e., $\mathcal{L}(v, v') = \mathbb{E}(T)$. Now, we are ready to prove Proposition 1.

**Proof of Proposition 1.** We will use a coupling argument by defining another stopping time, $\hat{T}$, which (a) stochastically dominates $T$, and (b) is easier to analyze. Then, we will be able to derive an upper bound on $\mathbb{E}(T)$ using $\mathbb{E}(\hat{T})$.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $U = \{U_i\}_{i \in \mathbb{N}}$ be a sequence of i.i.d. random variables with uniform distribution over $[0, 1]$ defined in this space. Then, construct the random variables $\{X_i\}_{i \in \mathbb{N}}$ as follows: for any $i$, let

$$X_i = \begin{cases} w, & \text{if } U_i > \varepsilon - \frac{\varepsilon}{|\mathcal{E}(X_{i-1})|}, \\ w', & \text{otherwise}, \end{cases}$$

where $w$ is the next vertex on the shortest path on $G$ under GSP from $X_{i-1}$ to $v'$, and $w' \in \mathcal{V}$ is any vertex on $G$ such that $(X_{i-1}, w') \in \mathcal{E}$. Note that $Z(t) = d(X_t, v')$ for any $t \geq 0$, as previously defined. However, we can simplify this process and construct an alternative sequence of random variables, $\{\hat{X}_i\}_{i \in \mathbb{N}}$, where for any $i$, instead of using the degree of the current vertex to determine the probability of success, we use the maximum degree of the graph, $\Delta$. Specifically, let

$$\hat{X}_i = \begin{cases} w, & \text{if } U_i > \varepsilon - \frac{\varepsilon}{\Delta}, \\ w', & \text{otherwise}. \end{cases}$$

Recall that $T = \min\{t \geq 0 : Z(t) = -d(v, v')\}$. Similarly, we let $\hat{T} = \min\{t \geq 0 : \hat{Z}(t) = -d(v, v')\}$ where $\hat{Z}(t) = d(\hat{X}_t, v')$ for any $t \geq 0$. By the construction, since $\Delta \geq |\mathcal{E}(v)|$ for any $v \in \mathcal{V}$, we have $\mathbb{P}(Z(t) \leq \hat{Z}(t), \forall t) = 1$. This implies that $T$ is stochastically dominated by $\hat{T}$ and thus $\mathbb{E}(T) \leq \mathbb{E}(\hat{T})$.

Finally, construct $\{Y_i\}_{i \in \mathbb{N}}$. For any $i$, let

$$Y_i = \begin{cases} -1, & \text{if } U_i > \varepsilon - \frac{\varepsilon}{\Delta}, \\ 1, & \text{otherwise}. \end{cases}$$

Note that $\hat{Z}(t) = \sum_{i=1}^t Y_i$ is a simple random walk with $p = 1 - \varepsilon + \frac{\varepsilon}{\Delta}$. Again by the construction, we can see that whenever $\hat{Z}(t)$ decreases by 1, i.e., $U_i > \varepsilon - \frac{\varepsilon}{\Delta}$, so does $\hat{Z}(t)$. On the other hand, when $U_i \leq \varepsilon - \frac{\varepsilon}{\Delta}$, $\hat{Z}(t)$ increases by 1 while the value of $\hat{Z}(t)$ may increase by 1, stay constant or decrease by 1. Therefore, we have $\mathbb{P}(\hat{Z}(t) \leq \hat{Z}(t), \forall t) = 1$.

As before, define $\tilde{T}_k = \min\{t \geq 0 : \hat{Z}(t) = -k\}$ with $k = d(v, v')$. Thus, the observation $\mathbb{P}(\hat{Z}(t) \leq \hat{Z}(t), \forall t) = 1$ implies that $\tilde{T}$ is stochastically dominated by $\tilde{T}_{d(v,v')}$ so that we have $\mathbb{E}(\tilde{T}) \leq \mathbb{E}\left(\tilde{T}_{d(v,v')}\right)$. By Lemma 4, we obtain $\mathcal{L}(v, v') = \mathbb{E}(T) \leq \mathbb{E}(\tilde{T}) < \mathbb{E}(\tilde{T}) \leq \frac{d_G(v, v')}{1 - 2\varepsilon(1 - \frac{\varepsilon}{\Delta})} \leq \frac{1}{1 - 2\varepsilon(1 - \frac{\varepsilon}{\Delta})} \leq \frac{1}{1 - 2\varepsilon}$. This completes the proof of Proposition 1.

**E.2 Proof of Proposition 2**

**Proof.** Let $\{X_t\}_{t \in \mathbb{N}}$ denote the noiseless trajectory and $\{\hat{X}_t\}_{t \in \mathbb{N}}$ the trajectory traversed using a GSP routine. Recall that under $\psi^{\text{SPC}}$, the agent informs the adversary whether she
is currently traversing an element of \( \{X_i\} \) or performing a GSP routine. Since the GSP routine is generated independently from the goal \( D \), for any \( v \in V \) we have

\[
P(D = v \mid \{X_i\}_{i \in \mathbb{N}}) = P(D = v \mid \{X_i\}_{1 \leq i \leq N}).
\]

Thus, from the perspective of the adversary, the game reduces to the noiseless case. Further, the strategy being trajectory revealing implies that the adversary may declare a prediction at \( t = 1 \) without loss of generality since he will gain no information by observing the future trajectory. Therefore, for any adversary strategy \( \chi \), \( q(\psi_0, \chi) = q(\psi^{\text{SPC}}, \chi) \) holds. Then, taking the supremum over all adversary strategies we conclude \( q^*(\psi_0) = q^*(\psi^{\text{SPC}}) \). \( \square \)

### E.3 Proof of Proposition 3

**Proof.** We will compute the expected intentional goal-hitting time under the assumption that the horizon is infinite. Then, when the horizon is finite, i.e., \( K < \infty \), we can simply truncate the goal-hitting time at time \( K \) which will result in a smaller expected delay.

For simplicity of notation, let us define \( q_t = P(T_{\text{IH}} = t) \) and note that \( q_0 = 0 \). In order to calculate the expected value of the intentional goal-hitting time, we condition on the event \( \{T_{\text{IH}} \geq t\} \) and obtain a recursive equation on \( q_t \):

\[
q_t = P(T_{\text{IH}} = t \mid T_{\text{IH}} \geq t)P(T_{\text{IH}} \geq t) + P(T_{\text{IH}} = t \mid T_{\text{IH}} < t)P(T_{\text{IH}} < t)
\]

\[
= P(T_{\text{IH}} = t \mid T_{\text{IH}} \geq t) \left( 1 - \sum_{i=1}^{t-1} q_i \right) + 0 \left( \sum_{i=1}^{t-1} q_i \right)
\]

\[
= p_{t-1}(1 - \varepsilon) \left( 1 - \sum_{i=1}^{t-1} q_i \right),
\]

where (a) follows from the following observation:

\[
P(T_{\text{IH}} = t \mid T_{\text{IH}} \geq t) = P(F_t = 1 \mid T_{\text{IH}} \geq t) = P(a_{t-1} = a^C, X_t = D, B_{t-1} = 1 \mid T_{\text{IH}} \geq t) = p_{t-1}(1 - \varepsilon).
\]

Solving recursively, we obtain:

\[
q_t = \begin{cases} 
\bar{q}, & \text{if } 1 \leq t < t^*, \\
\varepsilon t - t^* + 1 (1 - \varepsilon) (1 - (t^* - 1)\bar{q}), & \text{otherwise}.
\end{cases}
\]

Finally, we compute the expected intentional goal-hitting time under \( \psi^w_q \):

\[
E(T_{\text{IH}}) = E(T_{\text{IH}} \mid T_{\text{IH}} < t^*)P(T_{\text{IH}} < t^*) + E(T_{\text{IH}} \mid T_{\text{IH}} \geq t^*)P(T_{\text{IH}} \geq t^*)
\]

\[
= \frac{t^* - 1 + 1}{2} ((t^* - 1)\bar{q}) + \left( t^* - 1 + \frac{1}{1 - \varepsilon} \right) (1 - (t^* - 1)\bar{q})
\]

\[
= \frac{\bar{q}}{2} \left( \frac{1}{\bar{q}} - \varepsilon \right) \left( \frac{1}{\bar{q}} - \frac{1}{1 - \varepsilon} \right) + \frac{1}{\bar{q}} \left( 1 - \bar{q} \left( \frac{1}{\bar{q}} - \frac{1}{1 - \varepsilon} \right) \right)
\]

\[
= \frac{1}{2\bar{q}} + \frac{\bar{q} \varepsilon}{2(1 - \varepsilon)^2}.
\]

When \( K < \infty \), we will have \( E(T_{\text{IH}}^w) \leq E(T_{\text{IH}}^\psi) \leq \frac{1}{2\bar{q}} + \frac{\bar{q} \varepsilon}{2(1 - \varepsilon)^2} \). This completes the proof of Proposition 3. \( \square \)

38
E.4 Proof of Proposition 4

Proof. Recall that the maximal prediction risk, \( q^*(\psi^{\text{wt}}) \), is given by the following expression

\[
\sup_{\chi} q(\psi^{\text{wt}}, \chi) = \mathbb{P}(T^{\text{wt}} > K_n) + \sup_{\chi} \mathbb{P}(\dot{D}_{U^{\psi^{\text{wt}}}, \chi} = D, U_{\psi^{\text{wt}}, \chi} \leq T^{\text{wt}}).
\]

First, note that the probability with which the adversary wins due to the agent failing to reach the goal by the end of the horizon diminishes as the length of the horizon increases, i.e., \( \lim_{n \to \infty} \mathbb{P}(T^{\text{wt}} > K_n) = 0 \). Specifically, for the agent strategy \( \psi^{\text{wt}} \), we have

\[
\mathbb{P}(T^{\text{wt}} > K_n) \leq \frac{\mathbb{E}(T^{\text{wt}})}{K_n} \leq \frac{w}{K_n},
\]

by the Markov’s Inequality and the definition of the delay budget \( w \). Then, for fixed \( \psi^{\text{wt}} \) and \( w \), letting \( K_n \to \infty \) allows us to conclude \( \lim_{n \to \infty} \mathbb{P}(T^{\text{wt}} > K_n) = 0 \) since \( \lim_{n \to \infty} \frac{w}{K_n} = 0 \).

Denote by \( \hat{T}_{\text{IH}} \) the optimal Bayes estimator for \( T_{\text{IH}} \). That is, for a given trajectory realization \( x \), let \( \hat{T}_{\text{IH}}(x) = \arg \max_{\chi \in \mathcal{X}} \mathbb{P}(T_{\text{IH}} = t \mid X = x) \). Using this definition, we will next analyze the probability that the adversary wins by predicting correctly. Let \( \mathcal{X} \) denote the set of all trajectories under horizon \( K_n \). We have:

\[
\sup_{\chi} \mathbb{P}(\dot{D}_{U^{\psi}, \chi} = D, U_{\psi, \chi} \leq T^{\text{wt}}) = \sum_{x \in \mathcal{X}} \max_{t \geq 1} \mathbb{P}(T^{\text{wt}} = t \mid X = x) \mathbb{P}(X = x)
\]

\[
= \sum_{x \in \mathcal{X}} \max_{t \geq 1} \left[ \mathbb{P}(T_{\text{IH}} = t \mid X = x) + \mathbb{P}(T_{\text{IH}}^{\text{wt}} = t \mid X = x) \right] \mathbb{P}(X = x)
\]

\[
= \left( \sum_{x \in \mathcal{X}} \max_{t \geq 1} \mathbb{P}(X = x, T_{\text{IH}}^{\text{wt}} = t) \right) + \mathbb{P}(T_{\text{IH}}^{\text{wt}} = T_{\text{IH}}^{\text{wt}})
\]

\[
= \mathbb{P}(T_{\text{IH}} = T_{\text{IH}}^{\text{wt}}) + \mathbb{P}(T_{\text{IH}}^{\text{wt}} = T_{\text{IH}}^{\text{wt}}),
\]

where (a) follows from the observation that for each \( t \geq 1 \),

\[
\mathbb{P}(T = t \mid X = x) = \mathbb{P}(T = t, T_{\text{IH}} = T \mid X = x) + \mathbb{P}(T = t, T_{\text{IH}} \neq T \mid X = x)
\]

\[
\leq \mathbb{P}(T_{\text{IH}} = t, T_{\text{IH}} = T \mid X = x) + \mathbb{P}(T_{\text{IH}} \neq T \mid X = x)
\]

\[
\leq \mathbb{P}(T_{\text{IH}} = t \mid X = x) + \mathbb{P}(T_{\text{IH}} \neq T \mid X = x),
\]

and (b) from the definition of \( \hat{T}_{\text{IH}} \) as the optimal Bayes estimator for \( T_{\text{IH}}^{\text{wt}} \).

The two terms on the right-hand side of Eq. (23) correspond to the adversary’s success in predicting the intentional hitting time, \( \mathbb{P}(X = x, T_{\text{IH}}^{\text{wt}} = t) \), and the probability that the intentional hitting time differs from the actual hitting time, \( \mathbb{P}(T_{\text{IH}}^{\text{wt}} \neq T_{\text{IH}}) \). The next two results will provide upper bounds on these two terms, respectively. Notably, Proposition 11 is the core technical result of this section, where we leverage a path counting argument (Lemma 6 in Appendix F). Propositions 11 and 12 will be proved shortly in Sections E.4.1 and E.4.2, respectively.

Proposition 11. Fix \( n, K \in \mathbb{N} \), and \( \bar{q} \in [0, 1] \). Let \( G = (\mathcal{V}, \mathcal{E}) \) be a complete graph with \( |\mathcal{V}| = n \). Under the Water-Filling Strategy, \( \psi^{\text{wt}}_{\bar{q}} \):

\[
\mathbb{P}(\hat{T}_{\text{IH}} = T_{\text{IH}}^{\text{wt}}) \leq \bar{q}.
\]

(24)
Proposition 12. Fix $n, K \in \mathbb{N}$, and $\bar{q} \in [0,1]$. Let $G = (V,E)$ be a complete graph with $|V| = n$. Under the Water-Filling Strategy, $\psi_{\bar{q}}^{\text{wf}}$:

$$
\mathbb{P}(T_{\text{IH}}^{\text{wf}} \neq T_{\text{IH}}) \leq 1 - \left(1 - \frac{1}{n}\right)^K.
$$

We now return to the proof of Proposition 4. Applying Propositions 11 and 12 to the two terms in Eq. (23), we obtain

$$
\sup_{\chi} \mathbb{P}(\hat{D}_{U,\chi} = D, U_{\psi,\chi} \leq T_{\text{wf}}^{\text{IH}}) \leq \mathbb{P}(\hat{T}_{\text{IH}} = T_{\text{IH}}^{\text{wf}}) + \mathbb{P}(T_{\text{IH}}^{\text{wf}} \neq T_{\text{IH}}) \leq \bar{q} + 1 - \left(1 - \frac{1}{n}\right)^K.
$$

Hence, we have that $q^*(\psi_{\bar{q}}^{\text{wf}}) \leq \bar{q} + \delta_n$, where $\delta_n = \frac{w}{K} + 1 - \left(1 - \frac{1}{n}\right)^K$, and $\lim_{n \to \infty} \delta_n = 0$. This completes the proof of Proposition 4. \qed

**E.4.1 Proof of Proposition 11**

**Proof.** Let $\mathcal{X}$ denote the set of all trajectories under horizon $K$. Recall that $\hat{T}_{\text{IH}}$ is the optimal Bayes estimator for $T_{\text{IH}}^{\text{wf}}$. That is, for a given trajectory realization $x \in \mathcal{X}$ we let $\hat{T}_{\text{IH}}(x) = \arg \max_{t \geq 1} \mathbb{P}(T_{\text{IH}}^{\text{wf}} = t \mid X = x)$. Then, we can write

$$
\mathbb{P}(\hat{T}_{\text{IH}} = T_{\text{IH}}^{\text{wf}}) = \sum_{x \in \mathcal{X}} \mathbb{P}(X = x, \hat{T}_{\text{IH}}(x) = T_{\text{IH}}^{\text{wf}})
$$

$$
= \sum_{x \in \mathcal{X}} \mathbb{P}(\hat{T}_{\text{IH}}(x) = T_{\text{IH}}^{\text{wf}} \mid X = x) \mathbb{P}(X = x)
$$

$$
= \sum_{x \in \mathcal{X}} \max_{t \geq 1} \mathbb{P}(T_{\text{IH}}^{\text{wf}} = t \mid X = x) \mathbb{P}(X = x)
$$

$$
= \sum_{x \in \mathcal{X}} \max_{t \geq 1} \mathbb{P}(X = x, T_{\text{IH}}^{\text{wf}} = t).
$$

(26)

Now, we will compute the joint probability, $\mathbb{P}(X = x, T_{\text{IH}}^{\text{wf}} = t)$. For this purpose, we will first fix a vertex $v \in V$. Then, we will condition on this fixed $v$ being the goal and compute $\mathbb{P}(X = x, T_{\text{IH}}^{\text{wf}} = t \mid D = v)$. Finally, we will let $v = x_1$ and complete the derivation of $\mathbb{P}(X = x, T_{\text{IH}}^{\text{wf}} = t)$.

To find $\mathbb{P}(X = x, T_{\text{IH}}^{\text{wf}} = t \mid D = v)$, we let $A_i$ denote the event $\{X_i = x_i\}$ for each $i \geq 1$. Recall the feedback sequence under the Water-Filling Strategy, $\{\bar{F}_i\}_{i=1}^K$, where we have $\bar{F}_i = 1$ at time $i$ if a successful goal-attempt has occurred, and 0 otherwise. Similarly, we define the events $\bar{A}_i$ as follows:

$$
\bar{A}_i = \begin{cases} 
\{\bar{F}_i = 1\} & \text{if } i = t, \\
\{\bar{F}_i = 0\} & \text{otherwise}.
\end{cases}
$$

40
We can then write
\[
P(X = x, T^\text{wf}_H = t \mid D = v) = P \left( \bigcap_{i=1}^K (A_i \cap \tilde{A}_i) \mid D = v \right) \\
= P \left( \bigcap_{i=2}^K (A_i \cap \tilde{A}_i) \mid A_1 \cap \tilde{A}_1, D = v \right) P \left( A_1 \cap \tilde{A}_1 \mid D = v \right) \\
\quad \overset{(a)}{=} \prod_{i=2}^K P \left( A_i \cap \tilde{A}_i \mid \bigcap_{j=1}^{i-1} A_j \cap \tilde{A}_j, D = v \right) P \left( A_1 \cap \tilde{A}_1 \mid D = v \right),
\]
where \((a)\) is obtained by recursively conditioning on the events \(A_i \cap \tilde{A}_i\) for each \(i\). We next examine \(P \left( A_i \cap \tilde{A}_i \mid \bigcap_{j=1}^{i-1} A_j \cap \tilde{A}_j, D = v \right)\) for each \(i\) and observe that one of three cases must hold.

**Case 1:** \(i > t\). Given a successful attempt has already occurred at \(t\), the agent will do a random walk under \(\psi^\text{wf}\) at every period \(i > t\). The Random meta action yields \(\bar{F}_i = 0\) so that we obtain
\[
P \left( A_i \cap \tilde{A}_i \mid \bigcap_{j=1}^{i-1} A_j \cap \tilde{A}_j, D = v \right) = \frac{1}{|\mathcal{E}(x_{i-1})|}.
\]

**Case 2:** \(i < t\). Under \(\psi^\text{wf}\), if there has not been a successful goal-attempt yet, the agent will choose the meta action \(\tilde{a}^G\) with probability \(p_i\) and \(\tilde{a}^R\) otherwise. Consequently, the probability that the agent does not have a successful attempt at period \(i\) and travels to \(x_i\) can be written as
\[
P \left( A_i \cap \tilde{A}_i \mid \bigcap_{j=1}^{i-1} A_j \cap \tilde{A}_j, D = v \right) = \frac{P(\bar{F}_i = 0)}{|\mathcal{E}(x_{i-1})|}.
\]

**Case 3:** \(i = t\). Given that a successful attempt has not occurred up to time \(t\), the agent can reach her goal intentionally at \(t\) if (a) \(x_t\) is equal to \(v\), and (b) \(\bar{F}_i = 1\). Therefore,
\[
P \left( A_t \cap \tilde{A}_t \mid \bigcap_{j=1}^{t-1} A_j \cap \tilde{A}_j, D = v \right) = I(x_t = v)P(\bar{F}_t = 1).
\]

Substituting these cases into Eq. (27), we obtain the following:
\[
P \left( X = x, T^\text{wf}_H = t \mid D = v \right) = \left( \prod_{i=1}^{t-1} \frac{P(\bar{F}_i = 0)}{|\mathcal{E}(x_{i-1})|} \right) \cdot \frac{1}{|\mathcal{E}(x_{t-1})|} \cdot I(x_t = v)P(\bar{F}_t = 1) \cdot \prod_{i=t+1}^K \frac{1}{|\mathcal{E}(x_{i-1})|} \\
= \left( \prod_{i \neq t} \frac{1}{|\mathcal{E}(x_{i-1})|} \right) \left[ \prod_{i=1}^{t-1} P(\bar{F}_i = 0) \right] \cdot I(x_t = v)P(\bar{F}_t = 1). 
\]

We now evoke a key technical result, Lemma 6 in Appendix F, to obtain an upper bound on the term \(\prod_{i \neq t} \frac{1}{|\mathcal{E}(x_{i-1})|}\). Substituting Eq. (57) from Lemma 6 into Eq. (28), we have that
\[
P \left( X = x, T^\text{wf}_H = t \mid D = v \right) \leq \Delta G P(X^{[K]} = x) \left[ \prod_{i=1}^{t-1} P(\bar{F}_i = 0) \right] \cdot I(x_t = v)P(\bar{F}_t = 1),
\]

where \(\Delta G\) denotes the expected value of the goal function.
where $\bar{\Sigma}_G = \max_{v \in V} |\mathcal{E}(v)|$.

Now, let $x$ be a trajectory such that $x_t = v$, since otherwise we have $P(X = x, T_{th}^{wf} = t \mid D = v) = 0$. For this trajectory, we get:

$$P(X = x, T_{th}^{wf} = t \mid D = v) \leq \bar{\Sigma}_G P(\tilde{X}^{[K]} = x) \left( \prod_{i=1}^{t-1} P(\tilde{F}_i = 0) \right) P(\tilde{F}_i = 1)$$

$$= \frac{a}{\bar{\Sigma}_G} P(\tilde{X}^{[K]} = x) P(T_{th}^{wf} = t),$$

(30)

where (a) is due to the observation that $\left( \bigcap_{i=1}^{t-1} \{\tilde{F}_i = 0\} \right) \cap \{\tilde{F}_i = 1\}$ is equivalent to $\{T_{th}^{wf} = t\}$.

Next, we use Eq. (30) to characterize $P(X = x, T_{th}^{wf} = t)$ and replace $v = x_t$. We get:

$$P(X = x, T_{th}^{wf} = t) = P(X = x, T_{th}^{wf} = t \mid D = x_t) P(D = x_t)$$

$$\leq \bar{\Sigma}_G P(\tilde{X}^{[K]} = x) P(T_{th}^{wf} = t) P(D = x_t)$$

$$= \frac{a}{\bar{\Sigma}_G} P(\tilde{X}^{[K]} = x) q P(D = x_t)$$

$$= \frac{a}{\bar{\Sigma}_G} P(\tilde{X}^{[K]} = x) \left( \frac{q}{n} \right),$$

(31)

where step (a) uses that $P(T_{th}^{wf} = t) \leq \bar{q}$ for all $t \geq 1$ by design. Finally, step (b) is obtained by recalling that the goal is drawn uniformly at random from $V$.

Since the right hand side of Eq. (31) is independent of $t$, we further have

$$\max_{t \geq 1} P(X = x, T_{th}^{wf} = t) \leq \bar{\Sigma}_G P(\tilde{X}^{[K]} = x) \left( \frac{q}{n} \right).$$

(32)

Now, we substitute Eq. (32) into Eq. (26) and obtain

$$P(T_{th}^{cfg} = T_{th}^{wf} = t = \sum_{x \in \mathcal{X}} \max_{t \geq 1} P(X = x, T_{th}^{wf} = t) \leq \sum_{x \in \mathcal{X}} \bar{\Sigma}_G P(\tilde{X}^{[K]} = x) \left( \frac{q}{n} \right) = \frac{a}{\bar{\Sigma}_G} \frac{q}{n}.$$  

Since $\mathcal{X}$ is the set of all trajectories of length $K$ that can be traversed on $G$, $\tilde{X}^{[K]}$ always takes values in $\mathcal{X}$. Thus, step (a) follows from the observation that $\sum_{x \in \mathcal{X}} P(\tilde{X}^{[K]} = x) = 1$.

Lastly, noting that $\bar{\Sigma}_G = n - 1$ for complete graphs, we conclude $P(T_{th}^{cfg} = T_{th}^{wf} = t) \leq \bar{q}$. This completes the proof. \qed

E.4.2 Proof of Proposition 12

Proof. Write $T^{cfg} = T$ and $T_{th}^{wf} = T_{th}$ for notational simplicity. To compute the probability that $T$ and $T_{th}$ are different under the Water-Filling Strategy, we first condition on the value of $T_{th}$:

$$P(T_{th} \neq T) = \sum_{h=1}^{K+1} P(T_{th} \neq T \mid T_{th} = h) P(T_{th} = h).$$

(33)
Recall that if $B_t = 1$, then the agent’s action in period $t$ has been successful whereas she has been sent to a random vertex, otherwise. For each $1 \leq t \leq K + 1$, define the random variable $A_t$ as follows:

$$A_t = \mathbb{I}\{X_t = D \text{ and } B_{t-1} = 0\},$$

so that $A_t$ denotes whether there has been a goal hit that is not intentional, at time $t$.

Note that under $\psi^{\text{WF}}$, at any period $t$, we can have $A_t = 1$ under one of the following two scenarios:

1. The agent chooses the Random meta action and is sent to $D$, or
2. She chooses the Goal-Attempt meta action, fails with $\varepsilon$ probability (i.e., $B_{t-1} = 0$), and nature sends her to $D$.

Since both the Random meta action and the intrinsic uncertainty sample a vertex uniformly at random, independently from all other sources of randomness in the game, $\{A_t\}_{t=1}^{K+1}$ are independent. Hence, under $\psi^{\text{WF}}$, for any $t$ we have:

$$P(A_t = 1) = \frac{p_{t-1}\varepsilon + (1 - p_{t-1})}{\mathcal{E}(X_{t-1})} \leq \frac{1}{\Delta_G},$$

(34)

where $\Delta_G = \min_{v \in V} |\mathcal{E}(v)|$.

With this notation, given $T_{ih} = h$, we will have $T_{ih} = T$ if for all $t < h$, $A_t$ is 0. We can then write $P(T_{ih} = T \mid T_{ih} = h)$ as follows:

$$P(T_{ih} = T \mid T_{ih} = h) = \mathbb{P}\left(\bigcap_{t=1}^{h-1} \{A_t = 0\}\right)$$

$$\overset{(a)}{=} \prod_{t=1}^{h-1} P(A_t = 0)$$

$$\overset{(b)}{=} \left(1 - \frac{1}{\Delta_G}\right)^{h-1},$$

(35)

where (a) is due to the independence of $\{A_t\}_{t=1}^{K+1}$ and (b) follows from Eq. (34).

Finally, we can substitute Eq. (35) into Eq. (33) and obtain:

$$P(T_{ih} \neq T) = \sum_{h=1}^{K+1} P(T_{ih} \neq T \mid T_{ih} = h)P(T_{ih} = h)$$

$$\leq \sum_{h=1}^{K+1} \left[1 - \left(1 - \frac{1}{\Delta_G}\right)^{h-1}\right] P(T_{ih} = h)$$

$$= \mathbb{E}\left(1 - \left(1 - \frac{1}{\Delta_G}\right)^{T_{ih}-1}\right)$$

$$\overset{(a)}{=} 1 - \left(1 - \frac{1}{\Delta_G}\right)^{T_{ih}-1}$$

$$\overset{(b)}{=} 1 - \left(1 - \frac{1}{\Delta_G}\right)^{K},$$

43
where (a) and (b) are due to the Jensen’s Inequality and \( E(T_{IH}) \leq K + 1 \), respectively. Finally, we substitute \( \Delta \gamma = n - 1 \) and obtain \( \mathbb{P}(T_{IH} \neq T) \leq 1 - \left( \frac{1}{n} \right)^K \). This completes the proof. \( \square \)

E.5 Proof of Lemma 1

Proof. First, let us calculate the probability that the difference between the two goal-hitting times is equal to \( d \in \{1, \ldots, K_n\} \). By conditioning on the value of \( T_{IH} \), we get

\[
\mathbb{P}(T_{IH} - T_{\psi} = d) = \sum_{h=1}^{K_n+1} \mathbb{P}(T_{IH} - T_{\psi} = d | T_{IH} = h)\mathbb{P}(T_{IH} = h) \\
\leq \sum_{h=1}^{K_n+1} \mathbb{P}(T_{\psi} = h - d | T_{IH} = h) \\
= \sum_{h=1}^{K_n+1} \left( 1 - \frac{1}{n} \right)^{h-d-1} \frac{1}{n} \\
= 1 - \left( 1 - \frac{1}{n} \right)^{K_n-d} \\
\leq 1 - \left( 1 - \frac{1}{n} \right)^{K_n}. \tag{36}
\]

Next, we write the expected difference between the two goal-hitting times as below and use Eq. (36):

\[
\mathbb{E}(T_{IH} - T_{\psi}) = \sum_{d=1}^{K_n} d\mathbb{P}(T_{IH} - T_{\psi} = d) \\
\leq \sum_{d=1}^{K_n} d \left( 1 - \left( 1 - \frac{1}{n} \right)^{K_n} \right) \\
= \left( 1 - \left( 1 - \frac{1}{n} \right)^{K_n} \right) \frac{K_n(K_n+1)}{2} \\
= \sigma_n.
\]

Thus, we have \( \mathbb{E}(T_{IH}^\psi) \leq \mathbb{E}(T_{\psi}) + \sigma_n \) and \( \lim_{n \to \infty} \sigma_n = 0 \), as \( K_n \propto n \). This completes the proof. \( \square \)

E.6 Proof of Lemma 2

Proof. For any fixed agent strategy \( \psi \), let

\[
t(\psi) \in \arg\max_{t \in \mathbb{N}} \mathbb{P}(T_{IH}^\psi = t).
\]

Consider the adversary strategy \( \bar{\chi} \) where the adversary makes a prediction at \( t(\psi) \) equal to the agent’s state at that period, i.e., \( \bar{D}_{t(\psi)} = X_{t(\psi)} \). Under \( \bar{\chi} \), the prediction risk of \( \psi \) is given by

\[
q(\psi, \bar{\chi}) = \mathbb{P}(T_{\psi} = t(\psi)).
\]
The success event \( \{ T = t(\psi) \} \) of the adversary can be written as the following union of events

\[
\{ T = t(\psi) \} = \{ T = t(\psi), T_{IH} > t(\psi) \} \cup \{ T = t(\psi), T_{IH} = t(\psi) \},
\]

where under the first event the goal-hitting time is due to a random hit to the goal and under the latter it is due to an intentional hit. Ignoring the first event we have

\[
\mathbb{P}(T = t(\psi)) \geq \mathbb{P}(T = t(\psi), T_{IH} = t(\psi)).
\]

Next, we condition on \( \{ T_{IH} = t(\psi) \} \) and obtain:

\[
\mathbb{P}(T = t(\psi), T_{IH} = t(\psi)) = \mathbb{P}(T = t(\psi) \mid T_{IH} = t(\psi)) \mathbb{P}(T_{IH} = t(\psi))
\]

\[
\overset{(a)}{=} \left( 1 - \frac{1}{n^\delta_n} \right)^{t(\psi) - 1} \tilde{q}_\psi
\]

\[
\overset{(b)}{\geq} \left( 1 - \frac{1}{n} \right)^{K_n} \tilde{q}_\psi,
\]

where (a) follows upon observing that conditional on \( \{ T_{IH} = t(\psi) \} \), for the goal-hitting time to coincide with the intentional goal-hitting time, it must be the case that there have been no random hits to the goal in the first \( t(\psi) - 1 \) periods. Further, (b) is due to \( t(\psi) \leq K_n + 1 \).

Finally, we let \( \delta_n = 1 - \left( 1 - \frac{1}{n} \right)^{K_n} \) and observe that \( \lim_{n \to \infty} \delta_n = 0 \) holds, since \( K_n < n \). Thus, we conclude \( q^\ast(\psi) \geq q(\psi, \chi) \geq (1 - \delta_n) \tilde{q}_\psi \).

\[\square\]

### E.7 Proof of Proposition 5

**Proof.** In order to prove that the Water-Filling Strategy solves the given optimization problem, we will first write the expected intentional goal-hitting time as follows,

\[\mathbb{E}(T_{IH}) = \mathbb{E}(T_{IH} \mid T_{IH} \geq t^\ast) \mathbb{P}(T_{IH} \geq t^\ast) + \mathbb{E}(T_{IH} \mid T_{IH} < t^\ast) \mathbb{P}(T_{IH} < t^\ast).\]

Conditional on the event \( \{ T_{IH} \geq t^\ast \} \), where \( t^\ast = \frac{1}{\tilde{q}} - \frac{1}{1 - \tilde{\epsilon}} \), the Water-Filling Strategy minimizes the conditional expectation \( \mathbb{E}(T_{IH} \mid T_{IH} \geq t^\ast) \). This is because the Water-Filling Strategy tries to reach the goal greedily at every period after \( t^\ast \), by setting \( p_t = 1 \). Indeed, given that \( T_{IH}^w \) is strictly greater than \( t^\ast - 1 \), the conditional expected value \( \mathbb{E}(T_{IH}^w \mid T_{IH}^w \geq t^\ast) \) needs to be at least \( t^\ast - 1 + \frac{1}{1 - \tilde{\epsilon}} \), where \( \frac{1}{1 - \tilde{\epsilon}} \) is the stochastic shortest path diameter of the complete graph. However, the expression \( t^\ast - 1 + \frac{1}{1 - \tilde{\epsilon}} \) is exactly the conditional expectation of \( T_{IH}^w \) under the same event,

\[\mathbb{E}(T_{IH}^w \mid T_{IH}^w \geq t^\ast) \geq t^\ast - 1 + \frac{1}{1 - \tilde{\epsilon}} = \mathbb{E}(T_{IH}^w \mid T_{IH}^w \geq t^\ast).
\]

Hence, the term \( \mathbb{E}(T_{IH} \mid T_{IH} \geq t^\ast) \) is minimized under the Water-Filling Strategy. Recall that under \( \psi^w_q \), we have \( \mathbb{P}(T_{IH}^w = t) = \tilde{q} \) for all \( t < t^\ast \). Thus, the probability that the intentional goal-hitting time is greater than \( t^\ast \), i.e., \( \mathbb{P}(T_{IH}^w \geq t^\ast) \), is minimized under the Water-Filling Strategy subject to the requirement that at any period we have \( \mathbb{P}(T_{IH}^w = t) \leq \tilde{q} \).
To see this, suppose that there exists some feasible strategy $\psi$ such that $\mathbb{P}(T_{\text{IH}}^\psi \geq t^*) < \mathbb{P}(T_{\text{IH}}^{\psi_{\bar{q}}} \geq t^*) = 1 - \bar{q}t^*$. Equivalently, for $\psi$ we have

$$1 - \mathbb{P}(T_{\text{IH}}^\psi \geq t^*) = \mathbb{P}(T_{\text{IH}}^\psi < t^*) > \bar{q}t^*.$$ 

On the other hand, feasibility requires $\max_{t \in \mathbb{N}} \mathbb{P}(T_{\text{IH}}^\psi = t) \leq \bar{q}$. Thus, if $\psi$ is feasible, we must have

$$\mathbb{P}(T_{\text{IH}}^\psi < t^*) = \mathbb{P}(T_{\text{IH}}^\psi \leq t^* - 1) \leq \bar{q}(t^* - 1),$$

since otherwise the condition $\max_{t \in \mathbb{N}} \mathbb{P}(T_{\text{IH}}^\psi = t) \leq \bar{q}$ is violated. However, we have reached a contradiction: $\mathbb{P}(T_{\text{IH}}^\psi < t^*) \leq \bar{q}(t^* - 1)$ and $\mathbb{P}(T_{\text{IH}}^{\psi_{\bar{q}}} < t^*) > \bar{q}t^*$. Thus, we conclude that there does not exist any such strategy and that $\psi_{\bar{q}}^{\text{wf}}$ minimizes $\mathbb{P}(T_{\text{IH}}^\psi \geq t^*)$ among all feasible strategies. Also, note that $\psi_{\bar{q}}^{\text{wf}}$ maximizes $\mathbb{P}(T_{\text{IH}}^\psi < t^*)$ since the two expressions sum up to 1.

We have so far established that $\psi_{\bar{q}}^{\text{wf}}$ minimizes $\mathbb{E}(T_{\text{IH}}^\psi \mid T_{\text{IH}}^\psi \geq t^*)\mathbb{P}(T_{\text{IH}}^\psi \geq t^*)$. Consequently, if there exists a strategy $\psi$ such that $\mathbb{E}(T_{\text{IH}}^\psi) < \mathbb{E}(T_{\text{IH}}^{\psi_{\bar{q}}^{\text{wf}}})$, for this strategy we must have:

$$\sum_{t=1}^{t^*-1} t \mathbb{P}(T_{\text{IH}}^\psi = t) < \sum_{t=1}^{t^*-1} t \mathbb{P}(T_{\text{IH}}^{\psi_{\bar{q}}^{\text{wf}}} = t),$$

while satisfying $\sum_{t=1}^{\infty} \mathbb{P}(T_{\text{IH}}^\psi = t) = 1$. We will now show that this is not possible.

First, note that having $\mathbb{P}(T_{\text{IH}}^\psi = t) > \mathbb{P}(T_{\text{IH}}^{\psi_{\bar{q}}^{\text{wf}}} = t)$ for some $t < t^*$ results in infeasibility since $\mathbb{P}(T_{\text{IH}}^{\psi_{\bar{q}}^{\text{wf}}} = t) = \bar{q}$. Then, for the strict inequality in Eq. (37) to hold, we must have $\mathbb{P}(T_{\text{IH}}^{\psi_{\bar{q}}^{\text{wf}}} = t) = \mathbb{P}(T_{\text{IH}}^{\psi_{\bar{q}}^{\text{wf}}} = \hat{t}) - \delta$ for some $\hat{t} < t^*$ and $\delta > 0$. Since $\mathbb{P}(T_{\text{IH}}^\psi < t^*)$ is maximal under $\psi_{\bar{q}}^{\text{wf}}$, there must exist a set of indices $\{t_1, ..., t_k\}$ such that $k \geq 1$, $t_i \geq t^*$ for all $i = 1, ..., k$ and the following holds,

$$\mathbb{P}(T_{\text{IH}}^\psi = t_1) + ... + \mathbb{P}(T_{\text{IH}}^\psi = t_k) = \mathbb{P}(T_{\text{IH}}^{\psi_{\bar{q}}^{\text{wf}}} = t_1) + ... + \mathbb{P}(T_{\text{IH}}^{\psi_{\bar{q}}^{\text{wf}}} = t_k) + \delta.$$

Without loss of generality, let $k = 1$. For the expectations, because $\hat{t} < t^* \leq t_1$, it follows that

$$\mathbb{E}(T_{\text{IH}}^\psi) = \mathbb{E}(T_{\text{IH}}^{\psi_{\bar{q}}^{\text{wf}}}) - \delta \hat{t} + \delta t_1 > \mathbb{E}(T_{\text{IH}}^{\psi_{\bar{q}}^{\text{wf}}}).$$

Accordingly, we can conclude that it is not possible to construct a strategy $\psi$ for which the expected intentional goal-hitting time is smaller than that under the Water-Filling Strategy without violating the constraint. For this reason, the Water-Filling Strategy solves the given optimization problem and we conclude the proof.

**E.8 Proof of Proposition 6**

Similar to the counter $C_{\text{eff}}^\psi$, define another counter $C_t^\psi$ such that for all $t \geq 1$, we have $C_{t+1}^\psi + C_t^\psi = t$. Note that $C_t^\psi$ records the total time spent outside $\mathcal{V}_D$ up to time $t$.

**Proof.** Let us first find an upper bound on the expected total excursion time, that is, expected total time the agent spends outside $\mathcal{V}_D$. For this purpose, we will use a coupling
argument to analyze an alternative system in which transitions to $\mathcal{V} \setminus \mathcal{V}_D$ from both $\mathcal{V}_D$ and $\mathcal{V} \setminus \mathcal{V}_D$ are more likely. Further, we will ignore the cases in which $D$ is hit while returning back to $\mathcal{V}_D$ from $\mathcal{V} \setminus \mathcal{V}_D$. Then, the expected intentional goal-hitting time in this alternative system will provide an upper bound for the original $T_{\text{HH}}$. Denote the number of vertices in each clique by $m = \frac{2}{\varepsilon}$.

Under $\psi^k$, in each period $i \geq 1$, the probability that the state will make a transition from $\mathcal{V}_D$ to $\mathcal{V} \setminus \mathcal{V}_D$ can be written as

$$P(X_{i+1} \notin \mathcal{V}_D \mid X_i \in \mathcal{V}_D) = \frac{(1 - p_C\text{rf}(1 - \varepsilon))(k - 1)}{m + k - 2} := s_i.$$

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $U = \{U_i\}_{i \in \mathbb{N}}$ be a sequence of i.i.d. random variables with uniform distribution over $[0, 1]$ defined in this space. Construct the sequence of independent Bernouilli random variables $\{C_i\}_{i \in \mathbb{N}}$ as follows: let

$$C_i = \begin{cases} 1, & \text{if } U_i < s_i, \\ 0, & \text{otherwise,} \end{cases}$$

where the event $\{C_i = 1\}$ denotes that there is a transition leaving $\mathcal{V}_D$ at time $i$. Then, we can represent the number of transitions leaving $\mathcal{V}_D$ up to time $t$ by $Y_t = \sum_{i=1}^{t} C_i$.

Similarly, construct another sequence of i.i.d. Bernouilli random variables $\{\tilde{C}_i\}_{i \in \mathbb{N}}$ such that

$$\tilde{C}_i = \begin{cases} 1, & \text{if } U_i < \frac{k-1}{m+k-2}, \\ 0, & \text{otherwise,} \end{cases}$$

and define $\tilde{Y}_t = \sum_{i=1}^{t} \tilde{C}_i$. Observe that whenever $Y_t$ increases by 1, so does $\tilde{Y}_t$ even though $Y_t$ might stay constant or increase whenever $\tilde{Y}_t$ increases by 1. This implies that $P(Y_t \leq \tilde{Y}_t, \forall t) = 1$ and the expected number of transitions leaving $\mathcal{V}_D$ up to time $t$, i.e., $E(Y_t)$, satisfies $E(Y_t) \leq E(\tilde{Y}_t)$.

Next, define $Z_j$ as the time the agent spends outside $\mathcal{V}_D$ when she leaves $\mathcal{V}_D$ for the $j^{th}$ time. We know that under $\psi^k$ in each period $i \geq 1$, the probability that the state will stay in $\mathcal{V} \setminus \mathcal{V}_D$ is

$$P(X_{i+1} \notin \mathcal{V}_D \mid X_i \notin \mathcal{V}_D) = \frac{\varepsilon(m + k - 3)}{m + k - 2}.$$

This implies $\{Z_j\}_{j \in \mathbb{N}}$ is a sequence of i.i.d. geometric random variables with success probability $1 - \varepsilon + \frac{\varepsilon}{m+k-2}$.

Finally, we combine these results where we write $H = T_{\text{HH}}^{\text{wf}}$ to simplify notation whenever $T_{\text{HH}}^{\text{wf}}$ is a subscript. The expected total excursion time under $\psi^k$ satisfies:

$$E(C_H^{\text{op}}) \leq E\left(\sum_{j=1}^{\tilde{Y}_n} Z_j\right) \leq E\left(\sum_{j=1}^{\tilde{Y}_n} Z_j\right)^{(a)} = E(\tilde{Y}_H)E(Z_1) = E\left(\sum_{i=1}^{\tilde{Y}_H} \tilde{C}_i\right)E(Z_1) \leq \frac{E(T_{\text{HH}}^{\text{wf}})E(\tilde{C}_1)}{1 - \varepsilon + \frac{\varepsilon}{m+k-2}},$$

where $(a)$ follows from Wald’s Identity upon observing that $\{Z_j\}_{j \in \mathbb{N}}$ is an i.i.d. sequence independent from the nonnegative integer valued random variable $\tilde{Y}_H$. Likewise, $(b)$ is due to Wald’s Identity as $\{C_i\}_{i \in \mathbb{N}}$ is an i.i.d. sequence independent from the nonnegative integer valued $T_{\text{HH}}^{\text{wf}}$. 

47
If we ignore the cases in which the goal is reached while going back to $\mathcal{V}_D$ from another clique, the value we obtain can only be greater than the actual expected intentional goal-hitting time. Then, with time being indexed by $C^\text{wf}$, the strategy $\psi^k$ evolves in an identical manner with $\psi^\text{ref}$ and the expected intentional goal-hitting time under $\psi^\text{ref}$ provides an upper bound for $\mathbb{E}\left(C^\text{wf}_{IH}\right)$. Thus, writing $\mathbb{E}(T^k_{IH}) = \mathbb{E}\left(C^\text{wf}_{IH}\right) + \mathbb{E}\left(C^\alpha_{IH}\right)$, we conclude

$$\mathbb{E}(T^k) \leq \mathbb{E}\left(C^\text{wf}_{IH}\right) + \mathbb{E}\left(C^\alpha_{IH}\right) \leq \mathbb{E}(T^k_{IH}) + \mathbb{E}(T^\text{ref}_{IH})\frac{1}{(1-\varepsilon)(m+k-2)+\varepsilon}.$$

This completes the proof of Proposition 7.

**E.9 Proof of Proposition 7**

**Proof.** Recall that the maximal prediction risk, $q^\ast(\psi^k)$, is given by the following expression

$$\sup_\chi q(\psi^k,\chi) = \mathbb{P}(T^k > K_n) + \sup_\chi \mathbb{P}(\hat{D}_{U_{\psi^k,\chi}} = D, U_{\psi^k,\chi} \leq T^k).$$

First, note that we have:

$$\mathbb{P}(T^k > K_n) \leq \frac{\mathbb{E}(T^k)}{K_n} \leq \frac{w}{K_n},$$

by the Markov’s Inequality and the definition of the delay budget $w$. Thus, for fixed $\psi^k$ and $w$, letting $n \to \infty$ allows us to conclude $\lim_{n \to \infty} \mathbb{P}(T^k > K_n) = 0$ since $\lim_{n \to \infty} \frac{w}{K_n} = 0$.

Next, let us analyze the probability that the adversary wins by predicting correctly. Suppose the agent announces the clique containing her goal, $\mathcal{V}_D$, to the adversary at the beginning of the game through $\Gamma_1$. Then, the adversary can simply ignore whenever the agent is outside $\mathcal{V}_D$. Consequently, from the perspective of the adversary, the game reduces to predicting $D$ of an agent using $\psi^\text{ref}$ on a complete graph with $\frac{n}{k}$ vertices. Hence, by Proposition 4 the maximal prediction risk satisfies

$$\sup_\chi \mathbb{P}(\hat{D}_{U_{\psi^k,\chi}} = D, U_{\psi^k,\chi} \leq T^k) \leq \sup_\chi \mathbb{P}(\hat{D}_{U_{\psi^k,\chi}} = D, U_{\psi^k,\chi} \leq T^\text{ref}) \leq \bar{q} + 1 - \left(1 - \frac{1}{p_n}\right)^{K_n}.$$

Finally, we let $\delta_n = 1 - \left(1 - \frac{1}{p_n}\right)^{K_n} + \frac{w}{K_n}$ and observe that $\lim_{n \to \infty} \delta_n = 0$ since $K_n \ll \sqrt{n}$. Hence, we conclude $q^\ast(\psi^k) \leq \bar{q} + \delta_n$. This completes the proof of Proposition 7.

**E.10 Proof of Proposition 8**

**Proof.** Recall that the maximal prediction risk is given by the expression

$$q^\ast(\psi^p) = \sup_\chi q(\psi^p,\chi) = \mathbb{P}(T^p > K) + \sup_\chi \mathbb{P}(\hat{D}_{U_{\psi^p,\chi}} = D, U_{\psi^p,\chi} \leq T^p).$$

First, note that the probability with which the adversary wins due to the agent failing to reach the goal by the end of the horizon diminishes as the length of the horizon increases, i.e., $\lim_{K \to \infty} \mathbb{P}(T^p > K) = 0$. To see this, note that for the agent strategy $\psi^p$, we have

$$\mathbb{P}(T^p > K) \leq \frac{\mathbb{E}(T^p)}{K} \leq \frac{w}{K},$$

48
by the Markov’s Inequality and the definition of the delay budget \( w \). Then, for fixed \( \psi^p \) and \( w \), letting \( f_K = \frac{w}{K} \) implies \( \lim_{K \to \infty} f_K = 0 \) and \( \lim_{K \to \infty} \mathbb{P}(T^p > K) = 0 \).

Recall that \( T^p_{IH} \) denotes the optimal Bayes estimator for \( T_{IH} \). That is, for a given trajectory realization \( x \), let \( T^p_{IH}(x) = \arg \max_{t \geq 1} \mathbb{P}(T_{IH} = t \mid X = x) \). Using this definition, we will next analyze the probability that the adversary wins by predicting correctly. Let \( \mathcal{X} \) be the set of all trajectories under horizon \( K \). Then, we can write

\[
\sup_{\chi} \mathbb{P}(\tilde{D}_{U,\psi,\chi} = D, U_{\psi,\chi} \leq T^p) = \sum_{x \in \mathcal{X}} \max_{t \geq 1} \mathbb{P}(T^p = t \mid X = x) \mathbb{P}(X = x)
\]

\[
\leq \sum_{x \in \mathcal{X}} \max_{t \geq 1} [\mathbb{P}(T^p_{IH} = t \mid X = x) + \mathbb{P}(T^p_{IH} = T^p \mid X = x)] \mathbb{P}(X = x)
\]

\[
= \left( \sum_{x \in \mathcal{X}} \max_{t \geq 1} \mathbb{P}(X = x, T^p_{IH} = t) \right) + \mathbb{P}(T^p_{IH} \neq T^p)
\]

\[
= \mathbb{P}(\tilde{T}_{IH} = T^p_{IH}) + \mathbb{P}(T^p_{IH} \neq T^p),
\]

where (a) follows from the observation that for each \( t \geq 1 \), the following holds

\[
\mathbb{P}(T^p = t \mid X = x) = \mathbb{P}(T^p = t, T^p_{IH} = T^p \mid X = x) + \mathbb{P}(T^p = t, T^p_{IH} = T^p \mid X = x)
\]

\[
\leq \mathbb{P}(T^p_{IH} = t, T^p_{IH} = T^p \mid X = x) + \mathbb{P}(T^p_{IH} = T^p \mid X = x)
\]

and (b) from the definition of \( \tilde{T}_{IH} \) as the optimal Bayes estimator for \( T^p_{IH} \).

We note that the two terms on the right-hand side of (38) correspond to the adversary’s success in predicting the intentional hitting time, \( \mathbb{P}(X = x, T^p_{IH} = t) \), and the probability that the intentional hitting time differs from the actual hitting time, \( \mathbb{P}(T^p_{IH} \neq T^p) \). The next two results will provide upper bounds for these two terms, respectively.

We first state Proposition 13 which will provide an upper bound on the success probability of the optimal Bayes estimator for \( T_{IH} \), denoted by \( \tilde{T}_{IH} \). Precisely, for a given trajectory realization \( x \), \( \tilde{T}_{IH}(x) \) is defined by \( \tilde{T}_{IH}(x) = \arg \max_{t \geq 1} \mathbb{P}(T_{IH} = t \mid X = x) \). Notably, Proposition 13 is the core technical result of this section, where we leverage a path counting argument (Lemma 6 in Appendix F). The proof, given in Appendix E.10.1, is based on the insight gained from the proof of Proposition 4 for complete graphs, though with additional richness due to the complexity of the topology.

**Proposition 13.** Fix graph size \( n \in \mathbb{N} \), horizon \( K \in \mathbb{N} \) and target risk level \( q \in [0, 1] \). Let \( G = (\mathcal{V}, \mathcal{E}) \) be a graph with \( |\mathcal{V}| = n \). Under the Random Walk Water-Filling Strategy, \( \psi^p \),

\[
\mathbb{P}(\tilde{T}_{IH} = T^p_{IH}) \leq \frac{\Delta_G}{n},
\]

where \( \Delta_G \) is the maximum degree of \( G \), i.e., \( \Delta_G = \max_{v \in \mathcal{V}} |\mathcal{E}(v)| \).

Next, we state Proposition 14 quantifying the probability that \( T^p_{IH} \) and \( T^p \) are different under \( \psi^p \). The proof relies on a trajectory based analysis of \( T \) and explicitly uses the design of the strategy \( \psi^p \). It is provided in Appendix E.10.2.
Proposition 14. Fix graph size $n \in \mathbb{N}$, horizon $K \in \mathbb{N}$ and target risk level $\bar{q} \in [0,1]$. Let $G = (\mathcal{V}, \mathcal{E})$ be a graph with $|\mathcal{V}| = n$. Under the Random Walk Water-Filling Strategy, $\psi^p_q$, \[\mathbb{P}(T^p_{IH} \neq T^p) \leq 1 - \left(1 - \frac{1}{\Delta_G}\right)^K,\] where $\Delta_G$ is the minimum degree of $G$, i.e., $\Delta_G = \min_{v \in \mathcal{V}}|\mathcal{E}(v)|$.

Now, we return to the proof of Proposition 8 and apply Propositions 13 and 14 to the terms in (38), respectively. Thus, we obtain \[\sup_{\chi} \mathbb{P}(\hat{D}_{U_{\psi,x}} = D, U_{\psi,x} \leq T^p) \leq \mathbb{P}(\hat{T}_{IH} = T^p_{IH}) + \mathbb{P}(T^p_{IH} \neq T^p) \leq \bar{q} \frac{\Delta_G}{n} + 1 - \left(1 - \frac{1}{\Delta_G}\right)^K,\] and, we conclude $q^*(\psi^p) \leq \bar{q} \frac{\Delta_G}{n} + 1 - \left(1 - \frac{1}{\Delta_G}\right)^K + f_K$. This completes the proof. \[\square\]

E.10.1 Proof of Proposition 13

Proof. Let $\mathcal{X}$ denote the set of all trajectories under horizon $K$. Recall that $\hat{T}_{IH}$ is the optimal Bayes estimator for $T^p_{IH}$. That is, for a given trajectory realization $x \in \mathcal{X}$ we let $\hat{T}_{IH}(x) = \arg \max_{t \geq 1} \mathbb{P}(T^p_{IH} = t | X = x)$. Then, we can write \[\mathbb{P}(\hat{T}_{IH} = T^p_{IH}) = \sum_{x \in \mathcal{X}} \mathbb{P}(X = x, \hat{T}_{IH}(x) = T^p_{IH})\] \[= \sum_{x \in \mathcal{X}} \mathbb{P}(\hat{T}_{IH}(x) = T^p_{IH} | X = x) \mathbb{P}(X = x)\] \[= \sum_{x \in \mathcal{X}} \max_{t \geq 1} \mathbb{P}(T^p_{IH} = t | X = x) \mathbb{P}(X = x)\] \[= \sum_{x \in \mathcal{X}} \max_{t \geq 1} \mathbb{P}(X = x, T^p_{IH} = t).\] (39) Now, we will compute the joint probability, $\mathbb{P}(X = x, T^p_{IH} = t)$. For this purpose, we will first fix a vertex $v \in \mathcal{V}$. Then, we will condition on this fixed $v$ being the goal and compute $\mathbb{P}(X = x, T^p_{IH} = t | D = v)$. Finally, we will let $v = x_t$ and complete the derivation of $\mathbb{P}(X = x, T^p_{IH} = t)$.

To find $\mathbb{P}(X = x, T^p_{IH} = t | D = v)$, we now define a sequence of indicators $\{J_i\}_{i=1}^K$ where for each $i$, we write $J_i = 1$ only if a successful goal-attempt occurs at time $i$ under the Random Walk Water-Filling Strategy. Similarly, for each $i \geq 1$, we let $A_i$ denote the event $\{X_i = x_i\}$ and define the events $\tilde{A}_i$ as follows: \[\tilde{A}_i = \begin{cases} \{J_i = 1\} & \text{if } i = t, \\ \{J_i = 0\} & \text{otherwise}. \end{cases}\]
We can then write
\[
\mathbb{P}(X = x, T^*_{\text{Hh}} = t \mid D = v) = \mathbb{P} \left( \bigcap_{i=1}^{K} (A_i \cap \hat{A}_i) \mid D = v \right) = \mathbb{P} \left( \bigcap_{i=2}^{K} (A_i \cap \hat{A}_i) \mid A_1 \cap \hat{A}_1, D = v \right) \mathbb{P} \left( A_1 \cap \hat{A}_1 \mid D = v \right)
\]
\[= \left( \prod_{i=2}^{K} \mathbb{P} \left( A_i \cap \hat{A}_i \mid \bigcap_{j=1}^{i-1} A_j \cap \hat{A}_j, D = v \right) \right) \mathbb{P} \left( A_1 \cap \hat{A}_1 \mid D = v \right), \tag{40} \]
where (a) is obtained by recursively conditioning on the events \(A_i \cap \hat{A}_i\) for each \(i\).

For every \(t \geq 1\), let \(C_i^v\) denote the number of times a vertex which is a neighbor of the vertex \(v\) appears in the trajectory \(x\) up to time \(t\), i.e., \(C_i^v = \sum_{j=1}^{t} \mathbb{1}(x_j \in \mathcal{E}(v))\). Further, recall the feedback sequence under the Water-Filling Strategy, \(\{F_i\}_{i=1}^{K}\), where we have \(F_i = 1\) at time \(i\) if a successful goal-attempt has occurred, and \(0\) otherwise. Using these definitions, we next examine \(\mathbb{P} \left( A_i \cap \hat{A}_i \mid \bigcap_{j=1}^{i-1} A_j \cap \hat{A}_j, D = v \right)\) for each \(i\) and observe that one of four cases must hold.

**Case 1:** \(i > t\). Given a successful attempt has already occurred at \(t\), the agent will do a random walk under \(\psi^p\) at every period \(i > t\). The Random meta action yields \(J_i = 0\) so that we obtain
\[
\mathbb{P} \left( A_i \cap \hat{A}_i \mid \bigcap_{j=1}^{i-1} A_j \cap \hat{A}_j, D = v \right) = \frac{1}{\left| \mathcal{E}(x_{i-1}) \right|}.
\]

**Case 2:** \(i < t\) and \(x_{i-1} \notin \mathcal{E}(v)\). Under \(\psi^p\), for each \(i < t\), the agent will choose the Random meta action if she is outside the neighborhood of the goal, \(v\). Thus, we will have
\[
\mathbb{P} \left( A_i \cap \hat{A}_i \mid \bigcap_{j=1}^{i-1} A_j \cap \hat{A}_j, D = v \right) = \frac{1}{\left| \mathcal{E}(x_{i-1}) \right|}.
\]

**Case 3:** \(i < t\) and \(x_{i-1} \in \mathcal{E}(v)\). If the agent is currently inside the neighborhood of the goal \(v\), and if there has not been a successful goal-attempt yet, the agent will choose the corresponding action under \(\psi^w\) for period \(C_i^v\). Consequently, the probability that the agent does not have a successful attempt at period \(i\) and travels to \(x_i\) can be written as
\[
\mathbb{P} \left( A_i \cap \hat{A}_i \mid \bigcap_{j=1}^{i-1} A_j \cap \hat{A}_j, D = v \right) = \frac{\mathbb{P}(\bar{F}_{C_i^v} = 0)}{\left| \mathcal{E}(x_{i-1}) \right|}.
\]

**Case 4:** \(i = t\). Given that a successful attempt has not occurred up to time \(t\), the agent can reach her goal intentionally at \(t\) if (a) she is in the neighborhood of \(v\) at \(t - 1\), (b) \(x_t\) is equal to \(v\), and (c) \(\bar{F}_{C_{t-1}^v} = 1\). Therefore,
\[
\mathbb{P} \left( A_t \cap \hat{A}_t \mid \bigcap_{j=1}^{t-1} A_j \cap \hat{A}_j, D = v \right) = \mathbb{1}(x_{i-1} \in \mathcal{E}(v)) \mathbb{1}(x_t = v) \mathbb{P}(\bar{F}_{C_{t-1}^v} = 1).
\]
Substituting these cases into (40), we obtain the following:

\[
\mathbb{P}(X = x, T^p_{I_{I_H}} = t \mid D = v) = \left( \prod_{i=1}^{t-1} \left( \mathbb{I}[x_{i-1} \notin \mathcal{E}(v)] + \mathbb{I}[x_{i-1} \in \mathcal{E}(v)] \mathbb{P}(F_{C^v_{t-1}+1} = 0) \right) \right) \cdot \mathbb{I}[x_{t-1} \in \mathcal{E}(v)] \mathbb{I}[x_t = v] \mathbb{P}(F_{C^v_{t-1}+1} = 1) \cdot \prod_{i=t+1}^{K} \frac{1}{\mathbb{P}(I(x_i))}.
\]

Substituting (57) from Lemma 6 into (41), we have that

\[
\mathbb{P}(X = x, T^p_{I_{I_H}} = t \mid D = v) \leq \Delta_G \mathbb{P}(\hat{X}[K] = x) \cdot \left( \prod_{i=1}^{t-1} \left( \mathbb{I}[x_{i-1} \notin \mathcal{E}(v)] + \mathbb{I}[x_{i-1} \in \mathcal{E}(v)] \mathbb{P}(F_{C^v_{t-1}+1} = 0) \right) \right) \cdot \mathbb{I}[x_{t-1} \in \mathcal{E}(v)] \mathbb{I}[x_t = v] \mathbb{P}(F_{C^v_{t-1}+1} = 1).
\]

Now, let \( x \) be a trajectory such that \( x_t = v \), since otherwise we will have \( \mathbb{P}(X = x, T^p_{I_{I_H}} = t \mid D = v) = 0 \). For this trajectory, observing that \( \mathbb{I}[x_{t-1} \in \mathcal{E}(v)] \mathbb{I}[x_t = v] = 1 \) holds, we get:

\[
= \Delta_G \mathbb{P}(\hat{X}[K] = x) \cdot \left( \prod_{i=1}^{t-1} \left( \mathbb{I}[x_{i-1} \notin \mathcal{E}(v)] + \mathbb{I}[x_{i-1} \in \mathcal{E}(v)] \mathbb{P}(F_{C^v_{t-1}+1} = 0) \right) \right) \mathbb{P}(F_{C^v_{t-1}+1} = 1)
\]

\[
\overset{\text{(a)}}{=} \Delta_G \mathbb{P}(\hat{X}[K] = x) \mathbb{P}(T^w_{I_{I_H}} = C^v_{t-1} + 1),
\]

where (a) follows after substituting the values of \( \mathbb{I}[x_{t-1} \in \mathcal{E}(v)] \) and \( \mathbb{I}[x_{t-1} \notin \mathcal{E}(v)] \). Step (b) is obtained by observing that the event \( \bigcap_{i=1}^{t-1} \{ \tilde{F}_i = 0 \} \cap \{ \tilde{F}_t = 1 \} \) is equivalent to \( \{ T^w_{I_{I_H}} = t \} \).

Next, we use (43) to characterize \( \mathbb{P}(X = x, T^p_{I_{I_H}} = t) \) and replace \( v = x_t \). We get:

\[
\mathbb{P}(X = x, T^p_{I_{I_H}} = t) = \mathbb{P}(X = x, T^p_{I_{I_H}} = t \mid D = x_t) \mathbb{P}(D = x_t)
\]

\[
\leq \Delta_G \mathbb{P}(\hat{X}[K] = x) \mathbb{P}(T^w_{I_{I_H}} = C^v_{t-1} + 1) \mathbb{P}(D = x_t)
\]

\[
\overset{\text{(a)}}{=} \Delta_G \mathbb{P}(\hat{X}[K] = x) \tilde{q} \mathbb{P}(D = x_t)
\]

\[
\overset{\text{(b)}}{=} \Delta_G \mathbb{P}(\hat{X}[K] = x) \left( \frac{q}{n} \right),
\]

52
where step (a) uses that \( P(T_{IH}^p = t) \leq \tilde{q} \) holds for all \( t \geq 1 \) by design. Finally, step (b) is obtained by recalling that the goal is drawn uniformly at random from \( V \).

Since the right hand side of (44) is independent of \( t \), we further have

\[
\max_{t \geq 1} P(\mathbf{X} = \mathbf{x}, T_{IH}^p = t) \leq \sum_{x \in \mathcal{X}} P(\hat{X}^{[K]} = x) \left( \frac{\tilde{q}}{n} \right).
\]

Now, we substitute (45) into (39) and obtain

\[
P(\hat{T}_{IH} = T_{IH}^p) \leq \sum_{x \in \mathcal{X}} \max_{t \geq 1} P(\mathbf{X} = \mathbf{x}, T_{IH}^p = t)
\leq \sum_{x \in \mathcal{X}} \sum_{x \in \chi} P(\hat{X}^{[K]} = x) \left( \frac{\tilde{q}}{n} \right)
\stackrel{(a)}{=} \frac{\Delta_G}{n}.
\]

Since \( \mathcal{X} \) is the set of all trajectories of length \( K \) that can be traversed on \( G \), \( \hat{X}^{[K]} \) always takes values in \( \mathcal{X} \). Thus, step (a) follows from the observation that \( \sum_{x \in \mathcal{X}} P(\hat{X}^{[K]} = x) = 1 \).

This concludes the proof of Proposition 13.

**E.10.2 Proof of Proposition 14**

Proof. Write \( T^p = T \) and \( T_{IH}^p = T_{IH} \) for notational simplicity. To compute the probability that \( T \) and \( T_{IH} \) are different under the Random Walk Water-Filling Strategy, we first condition on the value of \( T_{IH} \):

\[
P(T_{IH} \neq T) = \sum_{h=1}^{K+1} P(T_{IH} \neq T | T_{IH} = h) P(T_{IH} = h).
\]

Recall that if \( B_t = 1 \), then the agent’s action in period \( t \) has been successful whereas she has been sent to a random vertex, otherwise. For each \( 1 \leq t \leq K + 1 \), define the random variable \( A_t \) as follows:

\[ A_t = I\{X_t = D \text{ and } B_{t-1} = 0\}, \]

so that \( A_t \) denotes whether there has been a goal hit that is not intentional, at time \( t \).

Note that under \( \psi^p \), at any period \( t \), we can have \( A_t = 1 \) under one of the following two scenarios:

1. the agent chooses the Random meta action and is sent to \( D \), or
2. she chooses the Goal-Attempt meta action, fails with \( \varepsilon \) probability (i.e., \( B_{t-1} = 0 \)), and nature sends her to \( D \).

Since both the Random meta action and the intrinsic uncertainty sample a vertex uniformly at random, independently from all other sources of randomness in the game, \( \{A_t\}_{t=1}^{K+1} \) are independent. Hence, under \( \psi^p \), for any \( t \) we have:

\[
P(A_t = 1) = \frac{\beta_{C_{t-1}} \varepsilon + (1 - \beta_{C_{t-1}})}{\left| \mathcal{E}(X_{t-1}) \right|} I\{X_{t-1} \in \mathcal{E}(D)\} \leq \frac{1}{\Delta_G},
\]

This concludes the proof of Proposition 13.
where \( C_t = \sum_{i=1}^t I\{X_i \in \mathcal{E}(D)\} \) and \( \Delta_G = \min_{v \in V} |\mathcal{E}(v)|. \)

With this notation, given \( T_{IH} = h \), we will have \( T_{IH} = T \) if for all \( t < h \), \( A_t \) is 0. We can then write \( P(T_{IH} = T \mid T_{IH} = h) \) as follows:

\[
P(T_{IH} = T \mid T_{IH} = h) = P\left( \bigcap_{t=1}^{h-1} \{A_t = 0\} \right)
\leq \prod_{t=1}^{h-1} P(A_t = 0)
\leq \left(1 - \frac{1}{\Delta_G}\right)^{h-1},
\]

where (a) is due to the independence of \( \{A_t\}_{t=1}^{K+1} \) and (b) follows from (47).

Finally, we substitute (48) into (46) and obtain:

\[
P(T_{IH} \neq T) = \sum_{h=1}^{K+1} P(T_{IH} \neq T \mid T_{IH} = h)P(T_{IH} = h)
\leq \sum_{h=1}^{K+1} \left[1 - \left(1 - \frac{1}{\Delta_G}\right)^{h-1}\right]E(T_{IH} = h)
= E\left(1 - \left(1 - \frac{1}{\Delta_G}\right)^{T_{IH}-1}\right)
\leq 1 - \left(1 - \frac{1}{\Delta_G}\right)^{E(T_{IH})-1}
\leq 1 - \left(1 - \frac{1}{\Delta_G}\right)^{K},
\]

where (a) and (b) are due to the Jensen’s Inequality and \( E(T_{IH}) \leq K + 1 \), respectively. This completes the proof of Proposition 14.

\[\text{Proof of Lemma 3}\]

To prove that a graph generated using the Erdős-Rényi random graph model belongs to the family \( G_n(p) \) with high probability, we will examine each property in Definition 6 separately. We will begin by verifying the upper bound on the degree using the Chernoff bound. Then, we will repeat similar arguments to prove the lower bound on the degree and the Neighborhood Overlap property.

**Regularity (Upper bound).** For each vertex \( v \in V \), define the event \( A_v = \{|\mathcal{E}(v)| \geq pn(1 + \gamma_n)\} \) to represent whether the degree of \( v \) is greater than \( pn(1 + \gamma_n) \). Since \( |\mathcal{E}(v)| \sim \text{Binomial}(n, p) \) under the Erdős-Rényi random graph model, using the Chernoff bound with \( \gamma_n \geq 0 \) we obtain

\[
P(A_v) = P(|\mathcal{E}(v)| \geq pn(1 + \gamma_n)) \leq \exp\left(\frac{-\gamma_n^2 pn}{2(1 + \gamma_n)}\right).
\]

54
Then, the probability that the maximum degree is at most $pn(1 + \gamma_n)$ can be written as

$$\mathbb{P}\left(\max_{v \in V} |\mathcal{E}(v)| \leq pn(1 + \gamma_n) \right) = 1 - \mathbb{P}\left(\bigcup_{v \in V} A_v \right)$$

$$\geq 1 - \sum_{v \in V} \mathbb{P}(A_v)$$

$$\geq 1 - n \exp\left(\frac{-\gamma_n^2 pn}{2} \right), \tag{50}$$

where (a) uses (49). Recall that $\gamma_n$ satisfies $\gamma_n \gg \sqrt{\log n/n}$ and $\lim_{n \to \infty} \gamma_n = 0$. This implies that $\exp(n\gamma_n^2) \gg n$ and consequently we obtain $\lim_{n \to \infty} n \exp\left(\frac{-\gamma_n^2 pn}{2 + \gamma_n} \right) = 0$. Hence, we conclude that the degrees of all vertices are upper-bounded by $pn(1 + \gamma_n)$ with high probability as $n \to \infty$, i.e., $\lim_{n \to \infty} \mathbb{P}\left(\max_{v \in V} \mathbb{P}(|\mathcal{E}(v)| \leq pn(1 + \gamma_n) \right) = 1$.

**Regularity (Lower bound).** As in the proof for the upper bound, for any vertex $v \in V$ define the event $B_v = \{ |\mathcal{E}(v)| \leq pn(1 - \gamma_n) \}$ to represent whether the degree of $v$ is less than $pn(1 - \gamma_n)$. Using the Chernoff bound with $\gamma_n \in [0, 1]$ we get

$$\mathbb{P}(B_v) = \mathbb{P}(|\mathcal{E}(v)| \leq pn(1 - \gamma_n)) \leq \exp\left(\frac{-\gamma_n^2 pn}{2} \right). \tag{51}$$

Similarly, we write the probability that the minimum degree is at least $pn(1 - \gamma_n)$ as

$$\mathbb{P}\left(\min_{v \in V} |\mathcal{E}(v)| \geq pn(1 - \gamma_n) \right) = 1 - \mathbb{P}\left(\bigcup_{v \in V} B_v \right)$$

$$\geq 1 - \sum_{v \in V} \mathbb{P}(B_v)$$

$$\geq 1 - n \exp\left(\frac{-\gamma_n^2 pn}{2} \right), \tag{52}$$

where (a) follows from (51).

Once again we recall $\gamma_n \gg \sqrt{\log n/n}$ and $\lim_{n \to \infty} \gamma_n = 0$. Then, we can write $\exp(n\gamma_n^2) \gg n$ and obtain $\lim_{n \to \infty} n \exp\left(\frac{-\gamma_n^2 pn}{2} \right) = 0$. Thus, we conclude that the degrees of all vertices are lower-bounded by $pn(1 - \gamma_n)$ with high probability as $n \to \infty$, i.e., $\lim_{n \to \infty} \mathbb{P}\left(\min_{v \in V} |\mathcal{E}(v)| \geq pn(1 - \gamma_n) \right) = 1$.

**Neighborhood Overlap.** For each pair of vertices $u, v \in V$, define the event $C_{uv} = \{ \frac{|\mathcal{E}(u) \cap \mathcal{E}(v)|}{|\mathcal{E}(v)|} \leq p(1 - \gamma_n) \}$. For simplicity of notation, denote the random variables $|\mathcal{E}(u) \cap \mathcal{E}(v)|$ and $|\mathcal{E}(v)|$ by $S_{uv}$ and $S_v$, respectively. Note that $S_v \sim \text{Binomial}(n, p)$, and $S_{uv} \sim \text{Binomial}(n, p^2)$. The probability that the Neighborhood Overlap property holds can be written as

$$\mathbb{P}\left(\frac{|\mathcal{E}(u) \cap \mathcal{E}(v)|}{|\mathcal{E}(v)|} \geq p(1 - \gamma_n), \forall u, v \in V \right) = 1 - \mathbb{P}\left(\bigcup_{u, v \in V} C_{uv}\right) \geq 1 - \sum_{u, v \in V} \mathbb{P}(C_{uv}). \tag{53}$$
Next, we will derive an upper bound on \( \mathbb{P}(C_{uv}) \). To this end, let \( \eta \in [0, 1] \) and for each vertex \( v \in V \), define the event \( D_v = \{ S_v < np(1 + \eta) \} \). Applying the Chernoff bound with \( \eta \geq 0 \) for \( S_v \), we obtain \( \mathbb{P}(D_v^c) = 1 - \mathbb{P}(D_v) \leq \exp \left( -\frac{\gamma_n^2 pm}{8 + 2 \gamma_n} \right) \).

Similarly, for each pair of vertices \( u, v \in V \), define the event \( D_{uv} = \{ S_{uv} > np^2(1 - \eta) \} \). Then, the Chernoff bound applied to \( S_{uv} \) with \( \eta \in [0, 1] \) implies \( \mathbb{P}(D_{uv}^c) = 1 - \mathbb{P}(D_{uv}) \leq \exp \left( -\frac{\gamma_n^2 np^2}{8} \right) \).

Now, observe that \( D_v \cap D_{uv} = \{ S_v < np(1 + \eta), S_{uv} > np^2(1 - \eta) \} \) holds so that we write \( \mathbb{P}(D_v \cap D_{uv}) \leq \mathbb{P} \left( \frac{S_{uv}}{S_v} > p \left( \frac{1-\eta}{1+\eta} \right) \right) \). Equivalently, we have

\[
\mathbb{P} \left( \frac{S_{uv}}{S_v} \leq p \left( \frac{1 - \eta}{1 + \eta} \right) \right) \leq 1 - \mathbb{P}(D_v \cap D_{uv}) = \mathbb{P}(D_v^c \cup D_{uv}^c) \leq \mathbb{P}(D_v^c) + \mathbb{P}(D_{uv}^c).
\]

Since \( \eta \in [0, 1] \), we further have

\[
\mathbb{P} \left( \frac{S_{uv}}{S_v} \leq p(1 - 2\eta) \right) \leq \mathbb{P} \left( \frac{S_{uv}}{S_v} \leq p \left( \frac{1 - \eta}{1 + \eta} \right) \right).
\]

Hence, choosing \( \eta = \frac{\gamma_n}{2} \), we obtain

\[
\mathbb{P}(C_{uv}) \leq \mathbb{P}(D_v^c) + \mathbb{P}(D_{uv}^c) \leq \exp \left( -\gamma_n^2 \frac{pm}{8 + 2 \gamma_n} \right) + \exp \left( -\gamma_n^2 \frac{np^2}{8} \right). \tag{54}
\]

Lastly, combining Eqs. (53) and (54) yields

\[
\mathbb{P} \left( \frac{|E(u) \cap E(v)|}{|E(v)|} \geq p(1 - \gamma_n), \forall u, v \in V \right) \geq 1 - \sum_{u,v \in V} \mathbb{P}(C_{uv}) \geq 1 - n^2 \left( \exp \left( -\gamma_n^2 \frac{pm}{8 + 2 \gamma_n} \right) + \exp \left( -\gamma_n^2 \frac{np^2}{8} \right) \right). \tag{55}
\]

We now observe that the assumptions \( \gamma_n \gg \sqrt{\log n/n} \) and \( \lim_{n \to \infty} \gamma_n = 0 \) imply \( \exp(n \gamma_n^2) \gg n^2 \). Consequently, we can conclude \( \lim_{n \to \infty} n^2 \exp \left( -\gamma_n^2 \frac{pm}{8 + 2 \gamma_n} \right) = 0 \) and \( \lim_{n \to \infty} n^2 \exp \left( -\gamma_n^2 \frac{np^2}{8} \right) = 0 \). Therefore, we obtain \( \lim_{n \to \infty} \mathbb{P} \left( \frac{|E(u) \cap E(v)|}{|E(v)|} \geq p(1 - \gamma_n), \forall u, v \in V \right) = 1 \) and conclude that Neighborhood Overlap property holds with high probability as \( n \to \infty \). Finally, we combine (50), (52), and (55) to obtain

\[
\mathbb{P}(G \in G_n) \geq 1 - \mathbb{P} \left( \bigcup_{v \in V} A_v + B_v \right) - \mathbb{P} \left( \bigcup_{u,v \in V} C_{uv} \right) \geq 1 - n \left( \exp \left( -\gamma_n^2 \frac{pm}{2 + \gamma_n} \right) + \exp \left( -\gamma_n^2 \frac{pm}{2} \right) \right) - n^2 \left( \exp \left( -\gamma_n^2 \frac{pm}{8 + 2 \gamma_n} \right) + \exp \left( -\gamma_n^2 \frac{np^2}{8} \right) \right) = 1 - \theta_n.
\]

We observe that \( \lim_{n \to \infty} \theta_n = 0 \) follows. This completes the proof that under the Erdős-Rényi model, Regularity and Neighborhood Overlap properties hold with high probability as \( n \to \infty \). \( \square \)
E.12 Proof of Proposition 9

Proof. The result follows from Proposition 8. Substituting \( \Delta_G = pn(1 - \gamma_n) \) and \( \Delta = pn(1 + \gamma_n) \), we obtain \( q^*(\psi^n) \leq \hat{q}p(1 + \gamma_n) + 1 - \left( 1 - \frac{1}{pn(1 - \gamma_n)} \right)^K + fK, \) where \( fK = \frac{w}{K} \). Since we have \( \lim_{n \to \infty} K_n = \infty \) and \( \lim_{K \to \infty} fK = 0 \), we get \( \lim_{n \to \infty} fK_n = 0 \). Next, we observe that \( \lim_{n \to \infty} 1 - \left( 1 - \frac{1}{pn(1 - \gamma_n)} \right)^K = 0 \) holds due to the assumption \( K_n \ll n \). Hence, we let \( \delta_n^* = 1 - \left( 1 - \frac{1}{pn(1 - \gamma_n)} \right)^K + fK_n \) and conclude the proof. \( \square \)

E.13 Proof of Proposition 10

Proof. To prove Proposition 10, we will derive the expected time between two consecutive trials and express the index of the successful trial in terms of \( T_{IH}^{ref} \). For any \( t \geq 1 \), define \( N_t = \sum_{i=1}^t I(X_i \in E(D)) \) to be the number of times the agent has been in the neighborhood of her goal up to time \( t \). Let \( U_n = \sup\{t \geq 1 : N_t < n\} \) denote the time of the \( n \)th occurrence of this event. Under \( \psi^n \), the system dynamics can equivalently be generated as follows:

1. Draw the goal, i.e., \( D \sim \text{Unif}(\mathcal{V}) \).
2. Draw the number of trials needed before the agent succeeds for the first time, \( R \), where the attempt probabilities are given by those under the Water-Filling Strategy on a complete graph with \( n \) vertices, \( \psi_q^n \). Note that \( R = T_{IH}^{ref} - 1 \).
3. Generate a random walk of length \( K \) on \( G, \{X_1, ..., X_K\} \). Then, set \( X_{K+1} = D \).
4. If \( U_R + 1 \leq K \), set \( X_{U_R+1, K+1} = D \) so that the agent indeed succeeds at her \( R \)th attempt, i.e., \( T_{IH} = U_R + 1 \). Replace \( \{X_{U_R+2, ..., X_K}\} \) with a random walk starting at \( D \).

Then, the observation \( T_{IH}^p = U_R + 1 \) implies

\[
\mathbb{E}(T_{IH}^p - 1) = \sum_{r=1}^\infty \mathbb{E}(U_r | R = r) \mathbb{P}(R = r) \overset{(a)}{=} \sum_{r=1}^\infty \mathbb{E}(U_r) \mathbb{P}(R = r), \tag{56}
\]

where (a) follows from the independence of \( U_r \) and \( R \).

Letting \( Y_i = U_i - U_{i-1} \) and \( Y_0 = 0 \), we can now write \( U_r = \sum_{i=1}^r Y_i \). To analyze \( \mathbb{E}(U_r) \), we state Lemma 5 whose proof is given in Appendix E.13.1.

**Lemma 5.** Let \( N_t = \sum_{i=1}^t \mathbb{I}(X_i \in E(D)) \) where \( \{X_t\}_{t \in \mathbb{N}} \) is the trajectory generated under \( \psi^n \). Define \( U_i = \sup\{t \geq 1 : N_t < i\} \) and \( Y_i = U_i - U_{i-1} \) for \( i \geq 1 \). Then, \( \mathbb{E}(Y_i) \leq \frac{1}{p(1 - \gamma_n)} \).

Using Eq. (56) and applying Lemma 5 in step (a), we obtain:

\[
\mathbb{E}(T_{IH}^p - 1) = \sum_{r=1}^\infty \mathbb{E} \left( \sum_{i=1}^r Y_i \right) \mathbb{P}(R = r) \overset{(a)}{=} \sum_{r=1}^\infty r \mathbb{P}(R = r) = \frac{\mathbb{E}(R)}{p(1 - \gamma_n)} \overset{(b)}{=} \frac{\mathbb{E}(T_{IH}^{ref}) - 1}{p(1 - \gamma_n)} \leq \frac{\mathbb{E}(T_{IH}^{ref})}{p(1 - \gamma_n)},
\]

where (b) is due to \( \mathbb{E}(R) = \mathbb{E}(T_{IH}^{ref}) - 1 \). Thus, we conclude the proof of Proposition 10. \( \square \)
E.13.1 Proof of Lemma 5

Proof. Let \( \tilde{Y} \) be a geometric random variable with success probability \( p(1 - \gamma_n) \), drawn independently from the rest of the system. We will now show that for all \( t \geq 1 \), \( \mathbb{P}(Y_1 > t) \leq \mathbb{P}(\tilde{Y} > t) \) holds so that \( Y_1 \) is stochastically dominated by \( \tilde{Y} \). This will give the upper bound on \( \mathbb{E}(Y_1) \).

First, note that \( Y_1 \) corresponds to the first hitting time of the random walk to the subset \( \mathcal{E}(D) \) and \( \tilde{Y}_i \) the \( i \)th return time to \( \mathcal{E}(D) \). We begin with \( t = 1 \). Write \( Z_t = \mathbb{I}(X_t \in \mathcal{E}(D)) \) for \( t \geq 1 \). Then,

\[
\mathbb{P}(Y_1 > 1) = \mathbb{P}(Z_1 = 0) = 1 - \frac{\left| \mathcal{E}(X_0) \cap \mathcal{E}(D) \right|}{\left| \mathcal{E}(X_0) \right|} \leq 1 - p(1 - \gamma_n) = \mathbb{P}(\tilde{Y} > 1),
\]

by the Neighborhood Overlap property of the family \( \mathcal{G}(p, \gamma_n) \). Next, suppose \( \mathbb{P}(Y_1 > t) \leq \mathbb{P}(\tilde{Y} > t) \) holds for some \( t \in \mathbb{N} \). Then, we have

\[
\mathbb{P}(Y_1 > t + 1) = \mathbb{P}(Y_1 > t) \mathbb{P}(Y_1 > t + 1 | Y_1 > t)
\]

\[
= \mathbb{P}(Y_1 > t) \mathbb{P}(Z_{t+1} = 0 | \{Z_i\}_{i=1}^{t} = 0)
\]

\[
= \mathbb{P}(Y_1 > t) \left( \sum_{v \in \mathcal{E}(X_{t-1})} \mathbb{P}(Z_{t+1} = 0 | \{Z_i\}_{i=1}^{t} = 0, X_t = v) \mathbb{P}(X_t = v | \{Z_i\}_{i=1}^{t} = 0) \right)
\]

\[
\overset{(a)}{=} \mathbb{P}(Y_1 > t) \left( \sum_{v \in \mathcal{E}(X_{t-1})} \mathbb{P}(Z_{t+1} = 0 | X_t = v) \mathbb{P}(X_t = v | \{Z_i\}_{i=1}^{t} = 0) \right)
\]

\[
\leq \mathbb{P}(Y_1 > t) \left( \max_{v \in \mathcal{E}(X_{t-1})} \mathbb{P}(Z_{t+1} = 0 | X_t = v) \right) \mathbb{P}(X_t = v | \{Z_i\}_{i=1}^{t} = 0)
\]

\[
\overset{(b)}{=} \mathbb{P}(Y_1 > t) \left( 1 - \frac{\left| \mathcal{E}(\tilde{v}) \cap \mathcal{E}(D) \right|}{\left| \mathcal{E}(\tilde{v}) \right|} \right)
\]

\[
\overset{(c)}{=} (1 - p(1 - \gamma_n))^{t+1}
\]

\[
= \mathbb{P}(\tilde{Y} > t + 1),
\]

where (a) is due to the Markov property, and (b) uses the definition \( \tilde{v} = \arg\max\{v \in \mathcal{E}(X_{t-1}) : \mathbb{P}(Z_{t+1} = 0 | X_t = v) \} \). Step (c) is again by the Neighborhood Overlap property and induction.

Since \( \mathbb{P}(Y_1 > t) \leq \mathbb{P}(\tilde{Y} > t) \) holds for all \( t \geq 1 \) as claimed, we obtain that \( Y_1 \) is stochastically dominated by \( \tilde{Y} \). Hence, we conclude \( \mathbb{E}(Y_1) \leq \mathbb{E}(\tilde{Y}) = (p(1 - \gamma_n))^{-1} \).

Finally, note that for \( Y_i \) with \( i > 1 \), the following still holds

\[
\mathbb{P}(Y_i > 1) = \mathbb{P}(Z_{U_i+1} = 0) = 1 - \frac{\left| \mathcal{E}(X_{U_i}) \cap \mathcal{E}(D) \right|}{\left| \mathcal{E}(X_{U_i}) \right|} \leq 1 - p(1 - \gamma_n) = \mathbb{P}(\tilde{Y} > 1),
\]

and a similar induction argument will conclude that \( \mathbb{P}(Y_i > t) \leq \mathbb{P}(\tilde{Y} > t) \) for all \( t \geq 1 \). Note that the only difference in the argument is in the first step since \( X_{U_i} \in \mathcal{E}(D) \) rather
than being chosen uniformly at random from \( \mathcal{V} \). Since the Neighborhood Overlap property of the family \( \mathcal{G}(p, \gamma_n) \) holds for all pairs of vertices, we can conclude \( \mathbb{E}(Y_i) \leq \mathbb{E}(\tilde{Y}) \) as before. This completes the proof. \( \square \)

### E.14 Proof of Theorem 3(b)

**Proof.** We prove Theorem 3(b) by combining Propositions 9 and 10. Fix \( n \in \mathbb{N} \), \( p \in [0,1] \) and let \( G = (\mathcal{V}, \mathcal{E}) \) be a graph from the family \( \mathcal{G}_n(p) \) such that \( |\mathcal{V}| = n \). Consider the Random Walk Water-Filling Strategy, \( \psi^\mathcal{V} \). By Proposition 10, expected delay under strategy \( \psi^\mathcal{V} \) satisfies \( \mathbb{E}(T^\mathcal{V}) \leq \frac{\mathbb{E}(T_\text{IH}^\mathcal{V})}{p(1-\gamma_n)} + 1 \). Then, to obtain \( \mathbb{E}(T^\mathcal{V}) \leq w \), it suffices to have \( \mathbb{E}(T_\text{IH}^\mathcal{V}) \leq \frac{p(w-1)}{2} \) because \( \gamma_n \in [0, \frac{1}{2}] \). Note that by Proposition 3, for \( \mathbb{E}(T_\text{IH}^\mathcal{V}) \leq y \), it suffices to set \( \bar{q} = \frac{1}{2y(1-c^\mathcal{V})^2} \) where \( c^\mathcal{V} = \frac{\epsilon}{(1-c^\mathcal{V})^2} \). Thus, letting \( y = \frac{p(w-1)}{2} \), and recalling \( Q(w) \leq q^b(\psi^\mathcal{V}) \), Proposition 9 implies

\[
Q(w) \leq pq(1 + \gamma_n) + \delta_n'
= \frac{p(1 + \gamma_n)}{p(w - 1) - 1 - c^\mathcal{V}} + \delta_n'
= \frac{1}{w - 1 - p^{-1} - c^\mathcal{V}p^{-1}} + \frac{\gamma_n}{p(w - 1) - 1 - c^\mathcal{V}p^{-1}} + \delta_n'
\leq \frac{1}{w - \lambda^\mathcal{V}(p)} + \frac{\gamma_n}{\sqrt{c^\mathcal{V} - c^\mathcal{V}p^{-1}}} + \delta_n',
\]

where \( \lambda^\mathcal{V}(p) = 1 + p^{-1} + c^\mathcal{V}p^{-1} \), \( \delta_n' = 1 - \left(1 - \frac{1}{pn(1-\gamma_n)}\right)^{K_n} + \frac{w}{K_n} \) and step (a) uses the condition \( w > \frac{2}{\lambda^\mathcal{V}(p)} + \frac{\sqrt{c^\mathcal{V} - c^\mathcal{V}p^{-1}}}{} + 1 \).

Define \( \delta_n = 1 + \frac{\gamma_n}{\sqrt{c^\mathcal{V} - c^\mathcal{V}p^{-1}}} + \delta_n' \). Recall that for the family \( \mathcal{G}_n(p) \) with \( \gamma = \gamma_n \), we have \( \lim_{n \to \infty} \gamma_n = 0 \). Therefore, the first term in \( \delta_n \) will approach 0 as \( n \to \infty \). Next, let us consider \( \delta_n' \). Since \( w \) is fixed and does not depend on \( n \), and since \( K_n \ll n \) by assumption, we also have \( \lim_{n \to \infty} \delta_n' = 0 \). Thus, we conclude \( \lim_{n \to \infty} \delta_n = 0 \). Finally, we get \( Q(w) \leq \frac{1}{w - \lambda^\mathcal{V}(p)} + \delta_n \) where \( \lim_{n \to \infty} \delta_n = 0 \). This completes the proof of Theorem 3. \( \square \)

### F Other Technical Results

Let \( G = (\mathcal{V}, \mathcal{E}) \) be a connected undirected graph. Fix \( x_0 \in \mathcal{V} \) and \( K \in \mathbb{N} \). Denote by \( \tilde{X}^{[K]} \) the (random) trajectory of a \( K \)-step random walk on \( G \) that starts from \( x_0 \). Recall that \( \Delta_G \) is the maximum degree of \( G \), i.e., \( \Delta_G = \max_{v \in \mathcal{V}} |\mathcal{E}(v)| \).

**Lemma 6.** Fix a \( K \)-step trajectory \( x \), and \( t \in \{1, 2, \ldots, K\} \). We have

\[
\prod_{i \in \{1, 2, \ldots, K\} \setminus \{t\}} \frac{1}{|\mathcal{E}(x_{i-1})|} \leq \Delta_G \mathcal{P} \left( \tilde{X}^{[K]} = x \right).
\]  

(57)
Proof. We have that
\[
\prod_{i \neq i} \frac{1}{|\mathcal{E}(x_{i-1})|} \leq \max_{1 \leq s \leq K} \prod_{i \neq i} \frac{1}{|\mathcal{E}(x_{i-1})|}
\]
\[
= \left( \max_{0 \leq s \leq K-1} |\mathcal{E}(x_s)| \right) \prod_{i=1}^{K} \frac{1}{|\mathcal{E}(x_{i-1})|}
\]
\[
\leq \left( \max_{v \in V} |\mathcal{E}(v)| \right) \prod_{i=1}^{K} \frac{1}{|\mathcal{E}(x_{i-1})|}
\]
\[
= \Delta G \prod_{i=1}^{K} \frac{1}{|\mathcal{E}(x_{i-1})|}
\]
\[
= \Delta G \mathbb{P}(\hat{X}^{[K]} = x),
\]
where the last step follows from the definition of a random walk.

\[\Box\]

G Generalization to Sparse Graphs

In this section, we again consider the family of graphs introduced in Section 8.2, \(\mathcal{F}(p, \gamma)\). However, instead of fixing the edge density \(p \in [0, 1]\) as in the family \(\mathcal{G}(p)\), we focus on a sequence of edge densities, \(\{p_n\}_{n \in \mathbb{N}}\) so that the graph families consist of more sparse graphs as \(n \to \infty\). Specifically, for each \(n \in \mathbb{N}\), we analyze graphs in the family \(\mathcal{F}(p_n, \gamma_n)\) where both \(p_n\) and \(\gamma_n\) depend on the graph size, \(n\). Definition 8 characterizes the family \(\mathcal{G}(\alpha)\) of graphs and lists the assumptions on the sequences \(\{\gamma_n\}_{n \in \mathbb{N}}\) and \(\{p_n\}_{n \in \mathbb{N}}\). We then prove Theorem 5 which generalizes Theorem 3 to sparse graphs and shows that additive overhead can be achieved for graphs in the family \(\mathcal{G}(\alpha)\).

Definition 8 (Family \(\mathcal{G}(\alpha)\) of Graphs). Fix \(\alpha \in (0, \frac{1}{2})\). Let \(\{p_n\}_{n \in \mathbb{N}}\) be a sequence such that \(p_n \in [0, 1]\) and \(p_n \gg n^{\alpha - \frac{1}{2}} \sqrt{\log n}\). For each \(n \in \mathbb{N}\), we define \(\mathcal{G}_n(\alpha) = \mathcal{F}(p_n, n^{-\alpha})\).

In the remainder of the section, we will focus on the family \(\mathcal{G}_n(\alpha)\) of graphs. Now, we are ready to state the main result of this section. The proof relies on the analysis of Theorem 3 and is given in Appendix G.1.

Theorem 5. Fix a noise level \(\varepsilon \in (0, 1)\) and \(\alpha \in (0, \frac{1}{2})\). Define sequences \(\{p_n\}_{n \in \mathbb{N}}, \{w_n\}_{n \in \mathbb{N}}, \{K_n\}_{n \in \mathbb{N}}\) and \(\{w_n\}_{n \in \mathbb{N}}\) such that:

1. \(p_n \gg n^{\alpha - \frac{1}{2}} \sqrt{\log n}\),
2. \(w_n = \frac{2}{p_n} \left( \frac{1}{2} + \frac{\sqrt{z}}{2} \right) + 1\),
3. \(K_n \gg w_n\) and \(K_n \ll p_n n\),
4. \(w_n \gg w_n\) and \(w_n \ll K_n\).

Then, there exists a sequence of families of graphs, \(\{\mathcal{G}_n(\alpha)\}_{n \in \mathbb{N}}\), such that the following are true.
(a) Suppose that \( \tilde{G} \) is a graph drawn from the Erdős-Rényi random graph model with \( n \) vertices and edge probability \( p_n \). Then,

\[ P(\tilde{G} \in \mathcal{G}_n(\alpha)) \geq 1 - \theta_n, \]

where \( \lim_{n \to \infty} \theta_n = 0 \).

(b) Fix \( n \in \mathbb{N} \), and suppose that \( G \in \mathcal{G}_n(\alpha) \). Then, the minimax prediction risk satisfies

\[ \frac{1}{2w_n + 1} \leq Q(w_n) \leq \frac{1}{2w_n - \lambda_n^\varepsilon} + \delta_n, \tag{58} \]

where \( \lambda_n^\varepsilon \) and \( \delta_n \) satisfy

\[ \lambda_n^\varepsilon \sim \frac{1 + \varepsilon^2}{(1 - \varepsilon)^2} \cdot \frac{1}{p_n}, \quad \text{and} \quad \delta_n \leq 1 \frac{K_n}{np_n} + \frac{w_n}{K_n}, \quad \text{as } n \to \infty. \]

Theorem 5 establishes that additive overhead can be achieved even for very sparse graphs generated from a random graph model, with average degree as low as \( p_n \sim \sqrt{n \log n} \).

Nevertheless, the upper bound that could be obtained by deliberately designing the network in Theorem 2 is stronger: it allows for a more sparse network (\( \tilde{p}_n \sim \sqrt{n} \)) and a smaller delay budget. The main reason behind this discrepancy is that the \( k \)-clique topology allows the agent to attempt her goal much more frequently. In contrast, while the approach based on a random walk delivers a similar prediction risk vs. delay tradeoff, it requires a larger delay budget to start with, since the agent is not able to attempt the goal as often.

The following corollary shows that, under a certain set of parameters, the first term in the upper bound in (58), \( \frac{1}{2w_n - \lambda_n^\varepsilon} \), dominates, and the upper bound can be further simplified as in Corollary 1. The result follows after observing that the conditions \( K_n \ll n^{1/3} \) and \( w_n \leq \sqrt{K_n} \) as \( n \to \infty \) imply \( \delta_n \ll \frac{1}{w_n} \). Consequently, we let \( \delta'_n = \delta_n w_n \), and conclude \( \lim_{n \to \infty} \delta'_n = 0 \).

**Corollary 1.** Suppose, in addition to the conditions in Theorem 5, that \( K_n \ll n^{1/3} \) and \( w_n \leq \sqrt{K_n} \) as \( n \to \infty \), then the minimax prediction risk satisfies

\[ Q(w_n) \leq \frac{1 + \delta'_n}{2w_n - \lambda_n^\varepsilon}, \quad n \in \mathbb{N}, \]

where \( \lambda_n^\varepsilon \) is as defined in (58), and \( \lim_{n \to \infty} \delta'_n = 0 \).

**G.1 Proof of Theorem 5**

**Proof.** First, we will verify that the assumptions of Theorem 3 continue to hold after substituting \( \gamma_n = n^{-\alpha} \) and prove (a). Then, we will characterize \( \lambda_n^\varepsilon \) and \( \delta_n \), using Propositions 9 and 10 to conclude the proof of Theorem 5(b).

Letting \( \gamma_n = n^{-\alpha} \), we observe that \( n^{-\alpha} \in [0, \frac{1}{2}] \) for any \( n \in \mathbb{N} \) and \( \lim_{n \to \infty} n^{-\alpha} = 0 \) as required. Also, note that we have \( n^{-\alpha} \gg \sqrt{\frac{\log n}{n}} \) since \( \alpha \in (0, \frac{1}{2}) \). This implies that graphs in \( \mathcal{G}_n(\alpha) \) satisfies the conditions of Definition 7. Therefore, for a graph \( \tilde{G} \) generated using
the Erdős-Rényi random graph model with edge probability \( p_n \), we can apply Lemma 3 to obtain \( P(\bar{G} \in \mathcal{G}_n(\alpha)) \geq 1 - \theta_n \), where

\[
\theta_n = n \left( \exp \left( \frac{-\gamma_n^2 p_n n}{2 + \gamma_n} \right) + \exp \left( \frac{-\gamma_n^2 p_n n}{2} \right) \right) - n^2 \left( \exp \left( \frac{-\gamma_n^2 p_n n}{8 + 2\gamma_n} \right) + \exp \left( \frac{-\gamma_n^2 n p_n^2}{8} \right) \right).
\]

By assumption (1) of Theorem 5, we obtain \( \lim_{n \to \infty} \theta_n = 0 \), as desired.

Next, we prove Theorem 5(b). By Proposition 10, expected delay under strategy \( \psi^p \) satisfies

\[
\mathbb{E}(T^p) \leq \frac{\mathbb{E}(T^w_{IH})}{p_n(1 - n^{-\alpha})} + 1.
\]

Then, to obtain \( \mathbb{E}(T^p) \leq w \), it suffices to have \( \mathbb{E}(T^w_{IH}) \leq p_n(w_n - 1)(1 - n^{-\alpha}) \). Note that by Proposition 3, for \( \mathbb{E}(T^w_{IH}) \leq y \), it suffices to set \( q = \frac{1}{2y - 1 - \varepsilon} \) where \( c^\varepsilon = \frac{\varepsilon}{(1 - \varepsilon)^2} \). Thus, letting \( y = p_n(w_n - 1)(1 - n^{-\alpha}) \), Proposition 9 implies

\[
Q(w_n) \leq p_n q (1 + n^{-\alpha}) + 1 - \left( 1 - \frac{1}{p_n n (1 - n^{-\alpha})} \right)^\kappa_n + \frac{w_n}{K_n}
= \frac{1 + n^{-\alpha}}{2w_n - \lambda_n^\varepsilon(p_n)} + \delta_n',
\]

where

\[
\lambda_n^\varepsilon(p_n) = 2 + 2n^{-\alpha}(w_n - 1) + \frac{1}{p_n} + \frac{c^\varepsilon}{p_n}, \quad \text{and}
\delta_n' = 1 - \left( 1 - \frac{1}{p_n n (1 - n^{-\alpha})} \right)^\kappa_n + \frac{w_n}{K_n}.
\]

We now analyze the first term on the right hand side of (59):

\[
\frac{1 + n^{-\alpha}}{2w_n - \lambda_n^\varepsilon(p_n)} = \frac{1}{[2w_n - \lambda_n^\varepsilon(p_n)] \cdot \left( \frac{1}{1 + n^{-\alpha}} \right)} \leq (a) \frac{1}{2w_n - \lambda_n^\varepsilon(p_n)} \cdot \left( 1 - n^{-\alpha} \right)
\leq (b) \frac{1}{2w_n - \lambda_n^\varepsilon(p_n) - 2w_n n^{-\alpha} + \lambda_n^\varepsilon(p_n) n^{-\alpha}}
\leq \frac{1}{2w_n - \left( 2n^{-\alpha}(w_n - 1) + 2 + \frac{1 + \varepsilon^2}{(1 - \varepsilon)^2 p_n} \right) - 2w_n n^{-\alpha} + \lambda_n^\varepsilon(p_n) n^{-\alpha}}
\]

For step (a) we use the Taylor expansion for \( \frac{1}{1+x} \) and observe \( \frac{1}{1+x} \geq 1 - x \), where \( x \) stands for \( x = n^{-\alpha} \). Step (b) follows from the following observation:

\[
\frac{1}{p_n} + \frac{\varepsilon}{(1 - \varepsilon)^2 p_n} \leq \frac{1 + \varepsilon^2}{(1 - \varepsilon)^2 p_n}.
\]

Then, we let \( \bar{\lambda}_n^\varepsilon = \left( 2n^{-\alpha}(w_n - 1) + 2 + \frac{1 + \varepsilon^2}{(1 - \varepsilon)^2 p_n} \right) + n^{-\alpha}(2w_n - \lambda_n^\varepsilon(p_n)) \). We note that both of the terms \( 2n^{-\alpha}(w_n - 1) \) and \( n^{-\alpha}(2w_n - \lambda_n^\varepsilon(p_n)) \) diminish to 0 as \( n \to \infty \) by the
Assumptions (1), (2) and (4) in Theorem 5. Hence, we conclude $\bar{\lambda}_n^{\epsilon} \sim \frac{1+\epsilon^2}{(1-\epsilon)^2} \cdot \frac{1}{n}$ as $n \to \infty$, as desired.

Having analyzed the first term in Eq. (59), we now turn to the analysis of $1 - \left(1 - \frac{1}{p_n n(1-n^{-\alpha})}\right)^{K_n}$.

To this end, let \( \{a_n\}_{n \in \mathbb{N}} \) and \( \{b_n\}_{n \in \mathbb{N}} \) be two sequences and use the following fact based on Taylor expansions: if \( \lim_{n \to \infty} b_n = \infty \) and \( \lim_{n \to \infty} a_n = 0 \), then

\[
1 - \left(1 - \frac{a_n}{b_n}\right)^{b_n} \leq 1 - a_n, \quad \text{as } n \to \infty.
\]

(60)

Since \( \lim_{n \to \infty} K_n = \infty \) and \( \lim_{n \to \infty} \frac{K_n}{p_n n(1-n^{-\alpha})} = 0 \) by Assumptions (1) and (3), we let \( a_n = \frac{K_n}{p_n n(1-n^{-\alpha})} \) and \( b_n = K_n \). Thus, we obtain:

\[
1 - \left(1 - \frac{1}{p_n n(1-n^{-\alpha})}\right)^{K_n} \leq 1 - \frac{K_n}{np_n(1-n^{-\alpha})} \leq 1 - \frac{K_n}{np_n},
\]

where the last relation is due to \( \lim_{n \to \infty} n^{-\alpha} = 0 \). Finally, we let \( \bar{\delta}_n = 1 - \left(1 - \frac{1}{p_n n(1-n^{-\alpha})}\right)^{K_n} + \frac{w_n}{K_n} \) and complete the proof by observing \( \bar{\delta}_n \leq 1 - \frac{K_n}{np_n} + \frac{w_n}{K_n} \) as $n \to \infty$. \( \Box \)