Data Complexity of Querying Description Logic Knowledge Bases under Cost-Based Semantics

Meghyn Bienvenu[1], **Quentin Manière**[2, 3]

[1]Université de Bordeaux, CNRS, Bordeaux INP, LaBRI, UMR 5800, Talence, France
[2]Department of Computer Science, Leipzig University, Germany
[3]Center for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI), Dresden/Leipzig, Germany
meghyn.bienvenu@labri.fr, quentin.maniere@uni-leipzig.de

Abstract

In this paper, we study the data complexity of querying inconsistent weighted description logic (DL) knowledge bases under recently-introduced cost-based semantics. In a nutshell, the idea is to assign each interpretation a cost based upon the weights of the violated axioms and assertions, and certain and possible query answers are determined by considering all (resp. some) interpretations having optimal or bounded cost. Whereas the initial study of cost-based semantics focused on DLs between $\mathcal{EL}_\bot$ and $\mathcal{ALCO}$, we consider DLs that may contain inverse roles and role inclusions, thus covering prominent DL-Lite dialects. Our data complexity analysis goes significantly beyond existing results by sharpening several lower bounds and pinpointing the precise complexity of optimal-cost certain answer semantics (no non-trivial upper bound was known). Moreover, while all existing results show the intractability of cost-based semantics, our most challenging and surprising result establishes that if we consider DL-Lite$^{\mathcal{H}}_{\text{bool}}$ ontologies and a fixed cost bound, certain answers for instance queries and possible answers for conjunctive queries can be computed using first-order rewriting and thus enjoy the lowest possible data complexity (AC^0).

1 Introduction

Ontology-mediated query answering (OMQA) has been extensively studied within the KR and database communities as a means of improving data access by exploiting semantic information provided by an ontology (Poggi et al. 2008; Bienvenu and Ortiz 2015; Xiao et al. 2018). Ontologies are typically formulated in decidable fragments of first-order logic (FO), with description logics (DLs) being a popular choice (Baader et al. 2017). Given an ontology (or TBox in DL parlance) $\mathcal{T}$, a dataset (or ABox) $\mathcal{A}$, and a query $q(\vec{x})$, the OMQA task boils down to finding the certain answers, i.e. tuples of constants $\vec{a}$ for which the instantiated query $q(\vec{a})$ is entailed from the knowledge base (KB) $(\mathcal{T}, \mathcal{A})$. Observe that if the input KB is inconsistent, every answer tuple is trivially a certain answer, so OMQA trivializes.

A prominent approach to tackling this issue is to adopt alternative inconsistency-tolerant semantics in order to be able to extract meaningful information from inconsistent KBs, cf. (Lembo et al. 2010) and surveys (Bienvenu and Bourgaux 2016; Bienvenu 2020). Many of these semantics are based upon repairs, defined as inclusion-maximal consistent subsets of the ABox. For example, the AR semantics considers the query answers that hold in all repairs, while the brave semantics returns those answers holding in at least one repair. Note that the line of work on repair-based semantics targets scenarios in which the TBox axioms are deemed fully reliable, so inconsistencies derive solely from errors in the ABox. However, in practice, it can be useful to allow for TBox axioms which typically hold but may admit rare exceptions. Such 'soft' ontology axioms can be addressed qualitatively, using generalized notions of repair that have been proposed for existential rule ontologies (Eiter, Lukasiewicz, and Predoiu 2016), or employing non-monotonic extensions of DLs that support defeasible axioms cf. (Bonatti, Lutz, and Wolter 2009; Giordano et al. 2013; Britz et al. 2021). Another option is to adopt a quantitative approach, using the recently proposed cost-based semantics for DL KBs (Bienvenu, Bourgaux, and Jean 2024a), henceforth referred to as (BBJ 2024) for succinctness.

In a nutshell, the idea is to annotate axioms and assertions with (possibly infinite) weights, which are used to assign a cost to each interpretation based upon the weights of the violated axioms and assertions (and taking into account also the number of violations of each TBox axiom). To query the KB, we may choose either to consider the set of interpretations achieving the optimal cost, or we may fix a cost bound k and consider all interpretations having cost at most k. We can then define the sets of certain and possible answers as those answers that hold respectively in all or some interpretation of optimal cost or bounded cost. As noted in (BBJ 2024), the optimal-cost certain answer semantics generalizes both the classical certain answer semantics and the AR semantics based upon weighted ABox repairs. Increasing the cost bound k beyond the optimal cost allows one to identify answers that are robust in the sense that they hold not only for the optimal-cost interpretations. Optimal- and bounded-cost possible answers generalize query satisfiability and can serve to compare candidate answers based upon their incompatibility with the KB.

The computational complexity of querying inconsistent weighted KBs under cost-based semantics was investigated in (BBJ 2024). Five central decision problems were considered: bounded-cost satisfiability (does there exist an inter-

Copyright © 2026, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

pretation with cost at most k?) plus query entailment under the four cost-based semantics. The complexity analysis was fairly comprehensive, considering both combined and data complexity, conjunctive and instance queries, and DLs ranging from the lightweight DL $\mathcal{EL}_\bot$ to the expressive DL $\mathcal{ALCO}$. One important question that was left open, however, was the data complexity of optimal-cost certain semantics (arguably the most useful of the semantics), for which no non-trivial upper bound was provided. Moreover, as the considered DLs allow neither inverse roles nor role inclusions, they do not yield any results for DLs of the DL-Lite family (Calvanese et al. 2007), which are the most commonly utilized in the context of OMQA.

The preceding considerations motivate us to embark on a more detailed data complexity analysis of cost-based semantics, considering various DL-Lite dialects and expressive DLs up to $\mathcal{ALCHIO}$. The results of our study are summarized in Table 1. A first major contribution, detailed in Section 4, is to provide a Δ_2^p upper bound for the optimal-cost certain and possible semantics, matching an existing lower bound for $\mathcal{EL}_\bot$ and a new lower bound we show for DL-Lite$_{\text{core}}$. This result is obtained by using an intricate quotient construction to establish a small interpretation property, which crucially does not depend on the considered cost. In Section 5, we strengthen a number of existing lower bounds for the bounded-cost semantics by showing that they hold even when for cost bound $k = 1$, as well as providing some new lower bounds for DL-Lite$_{\text{core}}$. Finally, our most challenging and surprising technical result (presented in Section 6) is to show that if we consider DL-Lite$_{\text{bool}}^{\mathcal{H}}$ ontologies and a fixed cost bound, then certain answers for instance queries and possible answers for conjunctive queries can be computed using first-order rewriting and thus enjoy the lowest possible data complexity (AC^0). Detailed proofs can be found in the appendix.

2 Preliminaries

We recall the syntax & semantics of description logics (DLs) and refer readers to (Baader et al. 2017) for further details.

Description logic knowledge bases We consider countably infinite sets $\mathsf{N_C}$, $\mathsf{N_R}$, and $\mathsf{N_I}$ of *concept names*, *role names*, and *individual names*. An *inverse role* has the form r^-, with $\mathsf{r} \in \mathsf{N_R}$. A *role* is either a role name or inverse role. We use $\mathsf{N_R^\pm} = \mathsf{N_R} \cup \{\mathsf{r}^- \mid \mathsf{r} \in \mathsf{N_R}\}$ for the set of roles. If $\mathsf{r} = \mathsf{s}^-$ is an inverse role, then r^- denotes s.

An $\mathcal{ALCIO}$ *concept* C is built according to the grammar $\mathsf{C}, \mathsf{D} ::= \top \mid \bot \mid \mathsf{A} \mid \{\mathsf{a}\} \mid \neg\mathsf{C} \mid \mathsf{C} \sqcap \mathsf{D} \mid \exists \mathsf{r}.\mathsf{D}$ where $\mathsf{A} \in \mathsf{N_C}$, $\mathsf{r} \in \mathsf{N_R^\pm}$, and $\mathsf{a} \in \mathsf{N_I}$. A concept $\{\mathsf{a}\}$ is called a *nominal*. An $\mathcal{ALCI}$ *concept* is a nominal-free $\mathcal{ALCIO}$ concept. An $\mathcal{EL}$ *concept* is an $\mathcal{ALCI}$ concept that uses neither negation, nor $\bot$, nor inverse roles ($\mathcal{EL}_\bot$ concepts may additionally use $\bot$).

An $\mathcal{ALCHIO}$ *TBox* is a finite set of *concept inclusions (CIs)* $\mathsf{C} \sqsubseteq \mathsf{D}$, where C, D are $\mathcal{ALCIO}$ concepts, and *role inclusions (RIs)* $\mathsf{r} \sqsubseteq \mathsf{s}$, where $\mathsf{r}, \mathsf{s} \in \mathsf{N_R^\pm}$. An $\mathcal{EL}$ *TBox* consists only of CIs between $\mathcal{EL}$ concepts. An *ABox* is a finite set of *concept assertions* $\mathsf{A}(\mathsf{a})$ and *role assertions* $\mathsf{r}(\mathsf{a}, \mathsf{b})$ where $\mathsf{A} \in \mathsf{N_C}$, $\mathsf{r} \in \mathsf{N_R}$, and $\mathsf{a}, \mathsf{b} \in \mathsf{N_I}$. We use $\mathsf{Ind}(\mathcal{A})$ for the set of individual names used in $\mathcal{A}$, and $\mathsf{N_C}(\mathcal{T})$ (resp. $\mathsf{N_R}(\mathcal{T})$) for the set of concept (resp. role) names used in $\mathcal{T}$. An $\mathcal{ALCHIO}$ *knowledge base (KB)* takes the form $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ with $\mathcal{T}$ an $\mathcal{ALCHIO}$ TBox and $\mathcal{A}$ an ABox.

We next introduce the syntax of some DLs of the DL-Lite family. A *basic concept* has the form A or $\exists\mathsf{r}$, with $\mathsf{A} \in \mathsf{N_C}$ and $\mathsf{r} \in \mathsf{N_R^\pm}$. A DL-Lite$_{\text{core}}^{\mathcal{H}}$ TBox is a finite set of CIs of the forms $\mathsf{C} \sqsubseteq \mathsf{D}$ and $\mathsf{C} \sqcap \mathsf{D} \sqsubseteq \bot$, with C, D basic concepts, and RIs $\mathsf{r} \sqsubseteq \mathsf{s}$, with $\mathsf{r}, \mathsf{s} \in \mathsf{N_R^\pm}$. We drop superscript $\cdot^{\mathcal{H}}$ if no role inclusions are admitted, and replace $\cdot_{\text{core}}$ by $\cdot_{\text{bool}}$ to indicate that C, D may be built from basic concepts using $\neg, \sqcap, \sqcup$.

The semantics of DL KBs is defined as usual in terms of interpretations $\mathcal{I} = (\Delta^\mathcal{I}, \cdot^\mathcal{I})$ with $\Delta^\mathcal{I}$ the non-empty *domain* and $\cdot^\mathcal{I}$ the *interpretation function*. An interpretation satisfies a CI $\mathsf{C} \sqsubseteq \mathsf{D}$ if $\mathsf{C}^\mathcal{I} \subseteq \mathsf{D}^\mathcal{I}$ and likewise for RIs. It satisfies an assertion $\mathsf{A}(\mathsf{a})$ if $\mathsf{a} \in \mathsf{A}^\mathcal{I}$ and $\mathsf{r}(\mathsf{a}, \mathsf{b})$ if $(\mathsf{a}, \mathsf{b}) \in \mathsf{r}^\mathcal{I}$. This notably requires ABox individuals to be interpreted as themselves, thus enforcing the *standard names assumption* (SNA) for the ABox individuals. The interpretation $\mathcal{I}_\mathcal{A}$ associated with an ABox $\mathcal{A}$ has domain $\mathsf{Ind}(\mathcal{A})$ and interprets concept and role names according to the assertions of $\mathcal{A}$:

$$\mathsf{A}^{\mathcal{I}_\mathcal{A}} := \{\mathsf{a} \mid \mathsf{A}(\mathsf{a}) \in \mathcal{A}\} \quad \mathsf{r}^{\mathcal{I}_\mathcal{A}} := \{(\mathsf{a}, \mathsf{b}) \mid \mathsf{r}(\mathsf{a}, \mathsf{b}) \in \mathcal{A}\}.$$

Given any interpretation $\mathcal{I}$, we use $\mathcal{I}|_\Delta$ to denote the restriction of $\mathcal{I}$ to a subdomain $\Delta \subseteq \Delta^\mathcal{I}$.

Queries First-order queries are given by formulas in first-order logic with equality. Since we wish to query DL KBs, we consider queries whose relational atoms can be either concept atoms $\mathsf{A}(t)$ or role atoms $\mathsf{r}(t, t')$, where $\mathsf{A} \in \mathsf{N_C}$, $\mathsf{r} \in \mathsf{N_R}$, and t, t' are *terms* (variables or individuals). We mostly focus on *conjunctive queries (CQs)* which have the form $\exists\vec{y}\psi$, where ψ is a conjunction of concept and role atoms, and $\vec{y}$ a tuple of variables from ψ. *Instances queries (IQs)* are CQs with a single atom. We also consider *acyclic* and *connected* CQs, meaning that their associated undirected graph (containing an edge $\{t, t'\}$ for each role atom $\mathsf{r}(t, t')$) has these properties. A Boolean CQ (BCQ) is a CQ that has no free variables. We write $\mathcal{I} \models q$ to indicate that an interpretation $\mathcal{I}$ satisfies a BCQ q.

Complexity classes For any syntactic object O such as a TBox, ABox, or query, we use $|O|$ to denote the *size* of O, meaning the encoding of O as a word over a suitable alphabet. Our complexity results concern the well-known complexity classes NP and coNP as well as Δ_2^p (deterministic polynomial time with access to an NP oracle) and AC^0. We omit the formal definition of AC^0, which is based upon circuits, as to understand our results, it suffices to know that it is in AC^0 (in data complexity) to test whether a Boolean first-order query is satisfied in a finite interpretation.

3 Cost-Based Semantics for Weighted KBs

In this section, we recall the definition of cost-based semantics from (BBJ 2024) and introduce the associated reasoning tasks. We also recall and prove some basic properties used in later sections.

Throughout the paper, we work with weighted KBs, whose assertions and axioms are annotated with weights:

	BCS^k, IQA_p^k, CQA_p^k	IQA_c^k	CQA_c^k	BCS, IQA_p, CQA_p	IQA_c, CQA_c	$\text{IQA}_{p,c}^{opt}$, $\text{CQA}_{p,c}^{opt}$
$\mathcal{EL}_\bot$ / $\mathcal{ALCHIO}$	NP ‡ Thm. 1, 3	coNP ‡ Thm. 1, 4	coNP ‡ Thm. 1, 4	NP † Thm. 1	coNP † Thm. 1	Δ_2^p † Thm. 2
$\text{DL-Lite}_{\text{core}}$ / $\text{DL-Lite}_{\text{bool}}^{\mathcal{H}}$	in AC^0 Thm. 8	in AC^0 Thm. 8	coNP Thm. 1, 7	NP Thm. 1, 5	coNP Thm. 1, 5	Δ_2^p Thm. 2, 6

Table 1: All results are completeness results, unless stated otherwise. †: lower bound from (BBJ 2024). ‡: lower bound for $k \geq 3$ from (BBJ 2024), improved to $k \geq 1$ in the present paper. For CQA, lower bounds already hold for connected acyclic BCQs. Results hold both for binary and unary encoding of weights, except for Δ_2^p-hardness (only for binary encoding).

Definition 1. *A* weighted knowledge base *(WKB)* $\mathcal{K}_\omega = (\mathcal{T}, \mathcal{A})_\omega$ *consists of a knowledge base* $(\mathcal{T}, \mathcal{A})$ *and a weight function* $\omega : \mathcal{T} \cup \mathcal{A} \mapsto \mathbb{N}_{>0} \cup \{\infty\}$. *We can similarly define* weighted TBoxes *($\mathcal{T}_\omega$) and* weighted ABoxes *($\mathcal{A}_\omega$).*

Intuitively, these weights can be viewed as the penalties incurred for violating assertions and axioms: those having cost 1 are the least reliable, while those assigned maximal weight ∞ should definitely be satisfied.

Interpretations will then be assigned costs based upon the sets of violations of the TBox axioms and ABox assertions. Note that differently from prior work, we also define violations of role inclusions, given by pairs of domain elements.

Definition 2. *Given an interpretation $\mathcal{I}$, the* set of violations of a concept inclusion $\mathsf{B} \sqsubseteq \mathsf{C}$ *in $\mathcal{I}$ is* $vio_{\mathsf{B} \sqsubseteq \mathsf{C}}(\mathcal{I}) = \mathsf{B}^\mathcal{I} \setminus \mathsf{C}^\mathcal{I}$, *the* set of violations of a role inclusion $\mathsf{r} \sqsubseteq \mathsf{s}$ *in $\mathcal{I}$ is* $vio_{\mathsf{r} \sqsubseteq \mathsf{s}}(\mathcal{I}) = \mathsf{r}^\mathcal{I} \setminus \mathsf{s}^\mathcal{I}$, *and the* violations of an ABox $\mathcal{A}$ *in $\mathcal{I}$ are* $vio_\mathcal{A}(\mathcal{I}) = \{\alpha \in \mathcal{A} \mid \mathcal{I} \not\models \alpha\}$.

Definition 3. *Let $\mathcal{K}_\omega = (\mathcal{T}, \mathcal{A})_\omega$ be a WKB. The* cost of an interpretation $\mathcal{I}$ *w.r.t. $\mathcal{K}_\omega$ is defined by:*

$$\omega(\mathcal{I}) = \sum_{\tau \in \mathcal{T}} \omega(\tau) |vio_\tau(\mathcal{I})| + \sum_{\alpha \in vio_\mathcal{A}(\mathcal{I})} \omega(\alpha)$$

The optimal cost *of $\mathcal{K}_\omega$ is $optc(\mathcal{K}_\omega) = \min_\mathcal{I}(\omega(\mathcal{I}))$. A WKB $\mathcal{K}_\omega$ is* k-satisfiable *if $\omega(\mathcal{I}) \leq k$ for some interpretation $\mathcal{I}$.*

We recall next the four cost-based semantics proposed in (BBJ 2024), which depend on whether one considers interpretations whose cost is less than a provided bound, or the interpretations having optimal cost, and whether the query is required to hold in all or at least one such interpretation.

Definition 4. *Let q be a BCQ, $\mathcal{K}_\omega = (\mathcal{T}, \mathcal{A})_\omega$ a WKB, and k an integer. We say that q is entailed by $\mathcal{K}_\omega$ under*

- k-cost bounded certain semantics*, written $\mathcal{K}_\omega \models_c^k q$, if $\mathcal{I} \models q$ for every interpretation $\mathcal{I}$ with $\omega(\mathcal{I}) \leq k$;*
- k-cost bounded possible semantics*, written $\mathcal{K}_\omega \models_p^k q$, if $\mathcal{I} \models q$ for some interpretation $\mathcal{I}$ with $\omega(\mathcal{I}) \leq k$;*
- opt-cost certain semantics*, written $\mathcal{K}_\omega \models_c^{opt} q$, if $\mathcal{I} \models q$ for every interpretation $\mathcal{I}$ with $\omega(\mathcal{I}) = optc(\mathcal{K}_\omega)$;*
- opt-cost possible semantics*, written $\mathcal{K}_\omega \models_p^{opt} q$, if $\mathcal{I} \models q$ for some interpretation $\mathcal{I}$ with $\omega(\mathcal{I}) = optc(\mathcal{K}_\omega)$.*

These semantics extend to non-Boolean CQs in the expected way, e.g. the opt-cost certain answers to a CQ $q(\vec{x})$ w.r.t. $(\mathcal{T}, \mathcal{A})_\omega$ are the tuples $\vec{a}$ from $\mathsf{Ind}(\mathcal{A})$ s.t. $\mathcal{K}_\omega \models_c^{opt} q(\vec{a})$.

If the underlying KB is satisfiable, then the certain and possible optimal-cost semantics coincide with query entailment and query satisfiability (or with classical notions of certain and possible answers, in the case of non-Boolean queries). These semantics are thus intended to be used when the underlying KB is inconsistent. The opt-cost certain answers identifies those answers that hold in the interpretations deemed most likely and have been shown to generalize previously considered weight-based repair semantics (BBJ 2024). By considering values of k beyond $optc(\mathcal{K}_\omega)$, we can use the k-cost bounded certain semantics to identify 'robust' answers which hold not only in the optimal-cost interpretations but also in those with close-to-optimal cost. By contrast, the opt-cost and k-cost bounded possible answers can serve to rank candidate answers based upon their degree of incompatibility with the WKB.

We now formalize the decision problems for cost-based semantics investigated in this paper, which differ depending on which cost bound is used and whether it is given as input:

- *Bounded cost satisfiability* (BCS) takes as input a WKB $\mathcal{K}_\omega = (\mathcal{T}, \mathcal{A})_\omega$ and an integer k and decides whether there exists an interpretation $\mathcal{I}$ with $\omega(\mathcal{I}) \leq k$.
- *k-cost satisfiability* (BCS^k) takes as input a WKB $\mathcal{K}_\omega = (\mathcal{T}, \mathcal{A})_\omega$ and decides whether there exists an interpretation $\mathcal{I}$ with $\omega(\mathcal{I}) \leq k$.
- *Bounded-cost certain (resp. possible) BCQ entailment* (CQA_c / CQA_p) takes as input a WKB $\mathcal{K}_\omega = (\mathcal{T}, \mathcal{A})_\omega$, a BCQ q and an integer k and decides whether $\mathcal{K}_\omega \models_c^k q$ (resp. $\mathcal{K}_\omega \models_p^k q$).
- *k-cost certain (resp. possible) BCQ entailment* (CQA_c^k / CQA_p^k) takes as input a WKB $\mathcal{K}_\omega = (\mathcal{T}, \mathcal{A})_\omega$ and a BCQ q and decides whether $\mathcal{K}_\omega \models_c^k q$ (resp. $\mathcal{K}_\omega \models_p^k q$).
- *Optimal-cost certain (resp. possible) BCQ entailment* (CQA_c^{opt} / CQA_p^{opt}) takes as input a WKB $\mathcal{K}_\omega = (\mathcal{T}, \mathcal{A})_\omega$ and a BCQ q and decides if $\mathcal{K}_\omega \models_c^{opt} q$ (resp. $\mathcal{K}_\omega \models_p^{opt} q$).

We will also consider the restrictions of the BCQ entailment problems to the case of instance queries, denoted by IQA_c, IQA_p, IQA_c^k, IQA_p^k, IQA_c^{opt} and IQA_p^{opt} respectively.

We shall study the *data complexity* of the preceding reasoning tasks. For the fixed-cost and optimal-cost decision problems, data complexity is measured with respect to the size of the input weighted ABox, while the size of the weighted TBox and query (if present) are treated as constants. For the bounded-cost problems, we measure complexity w.r.t. the weighted ABox and the input integer. Both the ABox weights and the input integer (if present) are assumed to be encoded in binary.

We conclude the section with some easy lemmas, which establish useful reductions between the decision problems.

Lemma 1. *BCS^k for* DL-Lite$_{\text{core}}$ *(resp. $\mathcal{EL}_\bot$) reduces to IQA$_p^k$ for* DL-Lite$_{\text{core}}$ *(resp. $\mathcal{EL}_\bot$). In particular, the same holds for BCS and IQA$_p$.*

Lemma 2. *IQA$_p^k$ for* DL-Lite$_{\text{core}}$ *(resp. $\mathcal{EL}_\bot$) reduces to the complement of IQA$_c^{k+1}$ for* DL-Lite$_{\text{core}}$ *(resp. $\mathcal{EL}_\bot$). In particular, IQA$_p$ reduces to the complement of IQA$_c$.*

Lemma 3. *For every $k \geq 0$, IQA$_p^k$ for* DL-Lite$_{\text{core}}$ *(resp. $\mathcal{EL}_\bot$) reduces to IQA$_p^{k+1}$ for* DL-Lite$_{\text{core}}$ *(resp. $\mathcal{EL}_\bot$) WKBs.*

4 An Upper Bound for the General Case

The aim of this section is to establish the next two theorems:

Theorem 1. *CQA$_p$ (resp. CQA$_c$) for $\mathcal{ALCHIO}$ is in* NP *(resp. in* coNP*).*

Theorem 2. *CQA$_p^{opt}$ and CQA$_c^{opt}$ for $\mathcal{ALCHIO}$ are in Δ_2^p.*

Starting from an interpretation $\mathcal{I}$ that satisfies (or does not satisfy) the query q of interest, the main technical ingredient is the construction of another interpretation $\mathcal{J}$ that behaves as $\mathcal{I}$ w.r.t. q, whose cost is at most the cost of $\mathcal{I}$, and whose domain has a size polynomially bounded by the size of the input ABox $\mathcal{A}$. This is fairly easy if query satisfaction is to be preserved, that is, for CQA$_p$: one can brutally collapse elements together according to their type using the well-known filtration technique (Baader et al. 2017).

Lemma 4. *Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})_\omega$ be an $\mathcal{ALCHIO}$ WKB, k an integer, and q a BCQ. If there exists an interpretation $\mathcal{I}$ such that $\omega(\mathcal{I}) \leq k$ and $\mathcal{I} \models q$, then there is an interpretation $\mathcal{J}$ such that $\omega(\mathcal{J}) \leq k$ and $\mathcal{J} \models q$, and whose domain $\Delta^{\mathcal{J}}$ has cardinality bounded polynomially in $|\mathcal{A}|$, with $|\mathcal{T}|$ and $|q|$ treated as constants, and independently from k.*

It becomes more challenging if query non-satisfaction is the property to preserve, that is, to address CQA$_c$. In particular, the solution adopted in (BBJ 2024, Proposition 8) for $\mathcal{ALCO}$ WKBs yields a polynomial bound that depends on the bounded cost k. This makes their technique adequate when the cost is fixed, that is, for CQA$_c^k$, or if the encoding of k is given in unary, but otherwise does not provide a polynomial upper bound w.r.t. data complexity. We deeply rework the approach, not only to support $\mathcal{ALCHIO}$ WKBs, but also to obtain a cost-independent bound as follows.

Lemma 5. *Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})_\omega$ be an $\mathcal{ALCHIO}$ WKB, k an integer, and q a BCQ. If there exists an interpretation $\mathcal{I}$ such that $\omega(\mathcal{I}) \leq k$ and $\mathcal{I} \not\models q$, then there is an interpretation $\mathcal{J}$ such that $\omega(\mathcal{J}) \leq k$ and $\mathcal{J} \not\models q$, and whose domain $\Delta^{\mathcal{J}}$ has cardinality that is bounded polynomially in $|\mathcal{A}|$, with $|\mathcal{T}|$ and $|q|$ treated as constants, and independently from k.*

With Lemmas 4 and 5, we obtain the NP and coNP upper bounds in Theorem 1 with standard guess-and-check procedures. For Theorem 2, the Δ_2^p algorithms proceed similarly but first identify the optimal cost via a binary search, using an exponential bound on the optimal cost (whenever finite).

Now, to prove Lemma 5, we rely on the adaptation of a quotient construction from (Manière 2022, Theorem 8) defined to answer *counting conjunctive queries* over $\mathcal{ALCHI}$ KBs. This construction was already reused by (BBJ 2024) for $\mathcal{ALCO}$ WKBs, and it is not too difficult to handle inverse roles and role inclusions by sticking closer to the original version. From the starting interpretation $\mathcal{I}$, our adaptation differs from theirs as it also takes as a parameter a subset $\mathcal{V} \subseteq \mathcal{T}$ of the considered TBox $\mathcal{T}$. In the constructed interpretation, we violate axioms of $\mathcal{V}$ exactly as in the original interpretation $\mathcal{I}$. Intuitively, one can think of these axioms in $\mathcal{V}$ as so expensive to violate that one cannot do better than in $\mathcal{I}$, while potential violations of axioms from $\mathcal{T} \setminus \mathcal{V}$ can be handled in a more systematic and structured manner. Violations of ABox assertions are easier to control and are preserved exactly as in the original interpretation $\mathcal{I}$. For an interpretation $\mathcal{J}$, we denote $vio_{\mathcal{V}}(\mathcal{J}) := \bigcup_{\tau \in \mathcal{V}} vio_\tau(\mathcal{J})$. Adapting the quotient technique yields the following:

Lemma 6. *Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ be an $\mathcal{ALCHIO}$ KB and q a BCQ. Let $\mathcal{V} \subseteq \mathcal{T}$ be a subset of $\mathcal{T}$ and $\mathcal{I}$ an interpretation such that $\mathcal{I} \not\models q$. There exists a polynomial p independent of $\mathcal{A}$, and an interpretation $\mathcal{J}$ satisfying the following:*

1. *$\mathcal{J} \not\models q$;*
2. *$vio_{\mathcal{A}}(\mathcal{J}) = vio_{\mathcal{A}}(\mathcal{I})$;*
3. *$\forall \tau \in \mathcal{V}, vio_\tau(\mathcal{J}) \subseteq vio_\tau(\mathcal{I})$;*
4. *$\left|\Delta^{\mathcal{J}}\right| \leq p(|\mathcal{A}| + |vio_{\mathcal{V}}(\mathcal{I})|)$.*

We now explain how to obtain Lemma 5, with an approach inspired from (Lutz and Manière 2024), where our Lemmas 5 and 6 respectively play the role of their Proposition 2 and Lemma 1. Let $\mathcal{K}_\omega = (\mathcal{T}, \mathcal{A})_\omega$ be an $\mathcal{ALCHIO}$ WKB, q a BCQ, k an integer, and $\mathcal{I}$ an interpretation such that $\mathcal{I} \not\models q$. For a given $\mathcal{V} \subseteq \mathcal{T}$, we use $\mathcal{J}_{\mathcal{V}}$ to denote the interpretation obtained by applying Lemma 6 with $\mathcal{V}$ the input set of axioms. We prove that there exists a subset $\mathcal{V} \subseteq \mathcal{T}$ such that (i) the size of $vio_{\mathcal{V}}(\mathcal{I})$ is bounded by a polynomial in $|\mathcal{A}|$ independent of k; and (ii) $\omega(\mathcal{J}_{\mathcal{V}}) \leq k$. To do so, we construct a sequence $\mathcal{V}_0 \subsetneq \mathcal{V}_1 \subsetneq \cdots \subsetneq \mathcal{V}_n \subseteq \mathcal{T}$ of $\mathcal{V}$'s that all satisfy item (i) and with $\mathcal{V}_n$ also satisfying item (ii). Note that $\mathcal{J}_{\mathcal{V}_n}$ is then the desired interpretation for Lemma 5: item (i) plus Point 4 from Lemma 6 gives the polynomial bound on the size of $\mathcal{J}_n$, while item (ii) and Point 1 in Lemma 6 ensure the desired properties w.r.t. the cost and query.

Initialization. Set $\mathcal{V}_0 := \emptyset$, which trivially satisfies item (i).

Induction step. Assume that, for some $i \geq 0$, we have successfully constructed $\mathcal{V}_i$ satisfying item (i); thus we have a polynomial p_i independent of k such that $|vio_{\mathcal{V}}(\mathcal{I})| \leq p_i(|\mathcal{A}|)$. If $\mathcal{V}_i$ also satisfies item (ii), then we are done. Otherwise $\omega(\mathcal{J}_{\mathcal{V}_i}) > k$, and since $\omega(\mathcal{I}) \leq k$, there exists an assertion or axiom τ from $\mathcal{K}$ that is violated at least once more in $\mathcal{J}_{\mathcal{V}_i}$ than in $\mathcal{I}$, *i.e.* $|vio_\tau(\mathcal{J}_{\mathcal{V}_i})| > |vio_\tau(\mathcal{I})|$. Note that due to Point 2 in Lemma 6, it is then clear that $\tau \notin \mathcal{A}$. Similarly, due to Point 3 in Lemma 6, we have $\tau \notin \mathcal{V}_i$. Therefore $\tau \in \mathcal{T} \setminus \mathcal{V}_i$. We set $\mathcal{V}_{i+1} := \mathcal{V}_i \cup \{\tau\}$. It remains to verify that $\mathcal{V}_{i+1}$ satisfies item (i). Note that $vio_{\mathcal{V}_{i+1}}(\mathcal{I}) = vio_{\mathcal{V}_i}(\mathcal{I}) \cup vio_\tau(\mathcal{I})$. The size of $vio_{\mathcal{V}_i}(\mathcal{I})$ is bounded adequately by $p_i(|\mathcal{A}|)$. For the size of $vio_\tau(\mathcal{I})$, recall that by choice of τ we have $|vio_\tau(\mathcal{I})| < |vio_\tau(\mathcal{J}_{\mathcal{V}_i})|$. We brutally bound $|vio_\tau(\mathcal{J}_{\mathcal{V}_i})|$ by $\left|\Delta^{\mathcal{J}_{\mathcal{V}_i}}\right|^2$. By Point 4 in Lemma 6, $\Delta^{\mathcal{J}_{\mathcal{V}_i}}$ has size bounded by $p(|\mathcal{A}| + |vio_{\mathcal{V}_i}(\mathcal{I})|)$, thus by $p(|\mathcal{A}| + p_i(|\mathcal{A}|))$. Overall, the size of $vio_{\mathcal{V}_{i+1}}(\mathcal{I})$ is bounded by $p_{i+1}(|\mathcal{A}|)$ where $p_{i+1}(x) := p_i(x) + (p(x + p_i(x)))^2$ is the desired polynomial independent of k.

Note that this procedure is guaranteed to terminate in at most $|\mathcal{T}|$ steps, which concludes the proof.

5 Lower Bounds

We first refine some existing lower bounds for the fixed-cost decision problems in $\mathcal{EL}_\bot$, then prove the lower bounds for DL-Lite listed in Table 1.

5.1 Lower Bounds in Extensions of $\mathcal{EL}_\bot$

We begin with two lower bounds showing that even if the cost k is fixed to 1, all considered reasoning tasks are NP-complete (or coNP-complete, depending on the task) already for $\mathcal{EL}_\bot$ WKBs. These results notably improve those from (BBJ 2024), where the cost was fixed to 3. Lemma 3 lifts our hardness proof to any fixed $k \geq 1$, and since the case of fixed $k = 0$ coincides with the usual semantics of $\mathcal{EL}$ KBs, the complexity w.r.t. fixed cost is now well understood.

Theorem 3. *For every $k \geq 1$, BCS^k, IQA$_p^k$ and CQA$_p^k$ for $\mathcal{EL}_\bot$ are* NP*-hard.*

Proof sketch. The reduction is from 3-SAT. Given a 3-CNF formula $\phi := \bigwedge_{i=1}^{\ell} \bigvee_{j=1}^{3} l_{i,j}$, where each $l_{i,j}$ is a literal over $v_1, \ldots, v_n$, we construct a WKB $(\mathcal{T}, \mathcal{A}_\phi)_\omega$. The ABox $\mathcal{A}_\phi$ contains $\mathsf{False}(\mathsf{a})$, $\mathsf{Bool}(\mathsf{a})$, and the additional assertions:

$$\mathsf{Var}(v_k) \text{ for } 1 \leq k \leq n \qquad \mathsf{clause}(\mathsf{a}, c_i) \text{ for } 1 \leq i \leq \ell$$
$$\mathsf{pos}_j(c_i, v_k) \text{ for } l_{i,j} = v_k \qquad \mathsf{neg}_j(c_i, v_k) \text{ for } l_{i,j} = \neg v_k$$

The TBox $\mathcal{T}$ has the following axioms:

$$\mathsf{Bool} \sqsubseteq \mathsf{True} \quad \mathsf{True} \sqcap \mathsf{False} \sqsubseteq \bot \quad \exists \mathsf{clause}.\mathsf{False} \sqsubseteq \mathsf{True}$$
$$\mathsf{Var} \sqsubseteq \exists \mathsf{val}.\mathsf{Bool} \quad \exists \mathsf{val}.\mathsf{True} \sqsubseteq \mathsf{True} \quad \exists \mathsf{val}.\mathsf{False} \sqsubseteq \mathsf{False}$$
$$\exists \mathsf{pos}_1.\mathsf{False} \sqcap \exists \mathsf{pos}_2.\mathsf{False} \sqcap \exists \mathsf{pos}_3.\mathsf{False} \sqsubseteq \mathsf{False}$$
(the six others combinations...)
$$\exists \mathsf{neg}_1.\mathsf{True} \sqcap \exists \mathsf{neg}_2.\mathsf{True} \sqcap \exists \mathsf{neg}_3.\mathsf{True} \sqsubseteq \mathsf{False}$$

The function ω assigns ∞ to all axioms and assertions, except for $\mathsf{Bool} \sqsubseteq \mathsf{True}$, which has weight 1. One can verify that ϕ is satisfiable iff $(\mathcal{T}, \mathcal{A}_\phi)_\omega$ is 1-satisfiable. □

For the case of IQA$_c^k$, we could use Lemma 2 to directly obtain a coNP-hardness proof for $k \geq 2$. We instead re-adapt the above proof to strengthen the result to every $k \geq 1$. Note that concept disjointness axioms are not even needed here.

Theorem 4. *For every $k \geq 1$, IQA$_c^k$ and CQA$_c^k$ for $\mathcal{EL}$ are* coNP*-hard.*

5.2 Lower Bounds in the DL-Lite Family

We now turn to the DL-Lite family, which inherits the upper bounds from Theorems 1 and 2. We begin with a proof that, when k is allowed to vary, all considered reasoning tasks are NP-hard (resp. coNP-hard) already for DL-Lite$_{\text{core}}$ WKBs.

Theorem 5. *BCS, IQA$_p$ and CQA$_p$ for* DL-Lite$_{\text{core}}$ *are* NP*-hard. IQA$_c$ and CQA$_c$ for* DL-Lite$_{\text{core}}$ *WKBs are* coNP*-hard.*

Note that, by virtue of Lemmas 1 and 2, it suffices to prove that BCS for DL-Lite$_{\text{core}}$ is NP-hard.

Proof sketch. We reduce from 3-COL, that is deciding whether a given graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is 3-colourable. All axioms of $\mathcal{T}$ are given infinite weight by ω and are as follows:

$$\exists s_i \sqcap \exists t_j \sqsubseteq \bot \text{ for } s, t \in \{\mathsf{r}, \mathsf{g}, \mathsf{b}\}, s \neq t, \text{ and } i, j \in \{1, 2\}$$
$$\exists s_1^- \sqcap \exists s_2^- \sqsubseteq \bot \text{ for } s \in \{\mathsf{r}, \mathsf{g}, \mathsf{b}\}$$

We choose an orientation $\mathcal{E}'$ of $\mathcal{E}$: for each $\{u, v\} \in \mathcal{E}$, we add either (u, v) or (v, u) in $\mathcal{E}'$. For each $e = (u, v) \in \mathcal{E}'$ and each $s \in \{\mathsf{r}, \mathsf{g}, \mathsf{b}\}$, we add $s_1(u, e)$ and $s_2(v, e)$ in the ABox $\mathcal{A}_\mathcal{G}$. Assertions in $\mathcal{A}_\mathcal{G}$ are given weight 1 by ω. It can be verified that $\mathcal{G} \in$ 3-COL iff $(\mathcal{T}, \mathcal{A}_\mathcal{G})_\omega$ is $4|\mathcal{E}|$-satisfiable. □

For optimal cost semantics, we establish a matching Δ_2^p lower bound. The proof adapts an existing construction from (Bourgaux 2016, Proposition 6.2.4) that establishes Δ_2^p-hardness of query entailment for DL-Lite KBs under preferred repair semantics, by reduction from deciding if a given variable is true in the lexicographically maximum truth assignment satisfying a given satisfiable CNF. We point out that, unlike the other lower bounds listed in Table 1, this result crucially relies upon a binary encoding of weights, intuitively because exponentially large weights are needed to perform lexicographic comparison of satisfying valuations.

Theorem 6. *IQA$_p^{opt}$, IQA$_c^{opt}$, CQA$_p^{opt}$, and CQA$_c^{opt}$ for* DL-Lite$_{\text{core}}$ *are Δ_2^p-hard.*

We now move to the case in which k is fixed. Notice indeed that the lower bound from Theorem 5 strongly relies on a varying k. For the certain semantics, we show that coNP-hardness holds even if k is fixed to 1, if we consider CQs. The proof is strongly inspired by the one of Theorem 4, especially with how to simulate the truth value assignment. The main difference is that we use acyclic CQs to circumvent the lack of nested concepts in DL-Lite$_{\text{core}}$.

Theorem 7. *For every $k \geq 1$, CQA$_c^k$ for* DL-Lite$_{\text{core}}$ *is* coNP*-hard. This holds already for connected acyclic BCQs and without concept disjointness axioms.*

6 Positive Results in the DL-Lite Family

In light of the lower bounds established in the previous section, we can only hope to achieve tractability for the DL-Lite family under a fixed cost (see Theorem 5). Furthermore, under the certain semantics, we established that CQA$_c^k$ answering is coNP-hard (Theorem 7) already for DL-Lite$_{\text{core}}$ and $k = 1$. This leaves us with two promising settings to explore: CQA$_p^k$ and IQA$_c^k$ for DL-Lite$_{\text{core}}$. This section establishes that both reasoning tasks enjoy the lowest possible complexity, that is, AC^0. Furthermore, we can even push this positive result to one of the most expressive logics of the DL-Lite family, namely DL-Lite$_{\text{bool}}^{\mathcal{H}}$.

Theorem 8. *For every integer $k \geq 1$, CQA$_p^k$ and IQA$_c^k$ for* DL-Lite$_{\text{bool}}^{\mathcal{H}}$ *are in* AC^0.

Our approach is based on first-order (FO) rewriting: given the weighted TBox $\mathcal{T}_{\omega_\mathcal{T}}$, query q and cost k, we construct an FO-query q' such that for every weighted ABox $\mathcal{A}_{\omega_\mathcal{A}}$, the following, here stated for CQA$_p^k$, holds:

$$(\mathcal{T}, \mathcal{A})_{\omega_\mathcal{T} \cup \omega_\mathcal{A}} \models_p^k q \quad \text{iff} \quad \mathcal{I}_{\mathcal{A}_{\omega_\mathcal{A}}} \models q'.$$

To make this formulation fully precise, we need to define the FO-interpretation $\mathcal{I}_{\mathcal{A}_{\omega_\mathcal{A}}}$ associated with a weighted ABox $\mathcal{A}_{\omega_\mathcal{A}}$, over which the rewritten query q' is evaluated. We argue that $\mathcal{A}_{\omega_\mathcal{A}}$ can be seen as a usual ABox augmented with extra assertions about the weights. We denote by $\mathcal{A}_k^{\omega_\mathcal{A}}$ the

extension of $\mathcal{A}$ with special concept and role assertions that encapsulate the relevant information about weights w.r.t. the fixed cost bound k: if a concept assertion $\omega_{\mathcal{A}}(\mathsf{A}(\mathsf{a})) = n$, then we add the assertion $\mathsf{W}^n_{\mathsf{A}}(\mathsf{a})$ if $n \leq k$, or assertion $\mathsf{W}^\infty_{\mathsf{A}}(\mathsf{a})$ if $n > k$. We proceed similarly for each role assertion $\mathsf{r}(\mathsf{a},\mathsf{b})$, adding respectively $\mathsf{w}^n_{\mathsf{r}}(\mathsf{a},\mathsf{b})$ or $\mathsf{w}^\infty_{\mathsf{r}}(\mathsf{a},\mathsf{b})$. Notice that we only need to introduce $(k+1)(\mathsf{N_C}(\mathcal{A})+\mathsf{N_R}(\mathcal{A}))$ fresh predicates and that computing $\mathcal{A}_k^{\omega_{\mathcal{A}}}$ from any reasonable representation of $\mathcal{A}_{\omega_{\mathcal{A}}}$ can be seen as a pre-processing step achieved by an AC^0 transducer. As $\mathcal{A}_k^{\omega_{\mathcal{A}}}$ is a usual ABox, its corresponding interpretation $\mathcal{I}_{\mathcal{A}_k^{\omega_{\mathcal{A}}}}$ is well defined and can serve as the desired interpretation $\mathcal{I}_{\mathcal{A}_{\omega_{\mathcal{A}}}}$.

Before explaining how to construct q', we sketch the main argument that allows for such a rewriting to exist and that ensures completeness of the claim (*i.e.* the $\Rightarrow$ direction in the formulation above). It relies on a (very!) small interpretation property: if an interpretation witnesses the desired behaviour w.r.t. the query and within the fixed cost, then there exists one that is completely trivial except on a small domain whose size is bounded by a *constant* w.r.t. the input weighted ABox. The different possibilities to interpret such a constant-size domain can thus all be encapsulated in the rewritten query q'. To facilitate the understanding of the rewriting, we first present this small interpretation property.

6.1 A (Very) Small Interpretation Property

Consider a WKB $\mathcal{K} = (\mathcal{T},\mathcal{A})_\omega$, a fixed cost k and a BCQ q. We introduce two distinct notions of types. The first is the usual one in DLs: a 1-*type* t is a subset of $\mathsf{N_C}(\mathcal{T}) \cup \{\exists \mathsf{r} \mid \mathsf{r} \in \mathsf{N}^\pm_{\mathsf{R}}(\mathcal{T})\}$. The 1-type of an element $e \in \Delta^{\mathcal{I}}$ is $\mathsf{tp}_{\mathcal{I}}(e) := \{\mathsf{A} \mid e \in \mathsf{A}^{\mathcal{I}}\} \cup \{\exists \mathsf{r} \mid e \in (\exists \mathsf{r})^{\mathcal{I}}, \mathsf{r} \in \mathsf{N}^\pm_{\mathsf{R}}\}$. This notion of 1-type captures the basic DL-Lite concepts, and thus, if two elements d and e have the same 1-type, then they violate exactly the same DL-Lite$_{\text{bool}}$ CIs.

The second notion of type is intended to capture the concepts that hold due to the ABox assertions. We define the *ABox type* $\mathsf{tp}_{\mathcal{A}}(\mathsf{a})$ of $\mathsf{a} \in \mathsf{Ind}(\mathcal{A})$ as:

$$\mathsf{tp}_{\mathcal{A}}(\mathsf{a}) := \mathsf{tp}_{\mathcal{I}_{\mathcal{A}}}(\mathsf{a}) \cup \{\exists^{>k}\mathsf{r} \mid \mathcal{A} \models \exists^{>k}\mathsf{r}, \mathsf{r} \in \mathsf{N}^\pm_{\mathsf{R}}(\mathcal{T})\},$$

where $\mathcal{A} \models \exists^{>k}\mathsf{r}$ means that there exists (at least) $k+1$ distinct individuals $\mathsf{b}_1,\dots,\mathsf{b}_{k+1} \in \mathsf{N_I}$ such that $\mathsf{r}(a,\mathsf{b}_1),\dots,\mathsf{r}(a,\mathsf{b}_{k+1}) \in \mathcal{A}$ (or $\mathsf{r}(\mathsf{b}_1,a),\dots,\mathsf{r}(\mathsf{b}_{k+1},a) \in \mathcal{A}$ if r is an inverse role). Notice that there are at most $2^{|\mathsf{N_C}(\mathcal{T})|+2|\mathsf{N_R}(\mathcal{T})|}$ possible 1-types, and at most $2^{|\mathsf{N_C}(\mathcal{T})|+4|\mathsf{N_R}(\mathcal{T})|}$ possible ABox types.

We write $\mathsf{r} \sqsubseteq_{\mathcal{T}} \mathsf{s}$ if (r,s) is in the transitive closure of $\{(\mathsf{p},\mathsf{p}) \mid \mathsf{p} \in \mathsf{N}^\pm_{\mathsf{R}}\} \cup \{(\mathsf{p},\mathsf{q}) \in \mathsf{N}^\pm_{\mathsf{R}} \times \mathsf{N}^\pm_{\mathsf{R}} \mid \mathsf{p} \sqsubseteq \mathsf{q} \in \mathcal{T}\}$. The following observation motivates the extension of 1-types into ABox types when a fixed cost k is considered:

Lemma 7. *Consider an interpretation $\mathcal{I}$ whose cost is $\leq k$. For every individual a and role $\mathsf{r} \in \mathsf{N}^\pm_{\mathsf{R}}$, if $\exists^{>k}\mathsf{r} \in \mathsf{tp}_{\mathcal{A}}(\mathsf{a})$, then $\exists \mathsf{s} \in \mathsf{tp}_{\mathcal{I}}(\mathsf{a})$ for every role $\mathsf{s} \in \mathsf{N}^\pm_{\mathsf{R}}$ such that $\mathsf{r} \sqsubseteq_{\mathcal{T}} \mathsf{s}$.*

We now prove that, if there exists an interpretation $\mathcal{I}$ whose cost is $\leq k$, then we find an interpretation $\mathcal{J}$ that behaves as $\mathcal{I}$ w.r.t. the query q, and whose cost is also $\leq k$, with all violations concentrated in a predictable portion of its domain. This portion of the domain of $\mathcal{J}$ is of course small, since the maximum number of violations is k, thus involving at most $2k$ distinct elements. By 'predictable', we mean that these violations take place among a small number of special individuals (constant number with respect to k) that can easily be identified, and on a small set of additional domain elements. To identify these special individuals, we rely on the following intuition: if an ABox type t is realized more than $2k$ times in $\mathcal{A}$, then there is a way to complete t without any 'local' violation. Indeed, if it was impossible to do so, then t being realized more than $2k$ times would always result in more than k violations and thus in a cost exceeding k, contradicting the very existence of $\mathcal{I}$. Therefore, only individuals with a rare ABox type may require a special treatment to keep the cost less than k. Formally, we say that an ABox type t is *rare* in $\mathcal{A}$ if $\#\{\mathsf{a} \in \mathsf{Ind}(\mathcal{A}) \mid \mathsf{tp}_{\mathcal{A}}(\mathsf{a}) = t\} \leq 2k$, and we use $\mathsf{RT}(\mathcal{A})$ for the set of rare ABox types in $\mathcal{A}$.

Now, when we start from an interpretation $\mathcal{I}$ with cost $\leq k$ and attempt to build $\mathcal{J}$, we preserve the interpretation of concepts and roles from $\mathcal{I}$ on those special individuals that have a rare ABox type. For $\mathcal{J}$ to behave like $\mathcal{I}$ with respect to the query q, we also preserve the interpretation on individuals occurring in q. We define the pre-core $\mathsf{pc}(\mathcal{A})$ as the set of individuals from $\mathcal{A}$ that have a rare ABox type, plus those query-related individuals, that is:

$$\mathsf{pc}(\mathcal{A}) := \{\mathsf{a} \mid \mathsf{tp}_{\mathcal{A}}(\mathsf{a}) \in \mathsf{RT}(\mathcal{A})\} \cup \mathsf{Ind}(q).$$

Unfortunately, the pre-core does not contain all the individuals that may be forced to participate in violations. As the following example illustrates, elements 'close' to the pre-core may also be forced to do so.

Example 1. *Consider the fixed cost $k := 3$ and the ABox $\mathcal{A} := \{\mathsf{A}(\mathsf{a}_0)\} \cup \bigcup_{i=0}^{7}\{\mathsf{r}(\mathsf{a}_i,\mathsf{b}_i),\mathsf{t}(\mathsf{b}_i,\mathsf{c}_i)\}$. The ABox type of a_0 is rare, others are not. Consider the TBox $\mathcal{T}$ with axioms:*

$$\mathsf{A} \sqsubseteq \exists \mathsf{u} \sqcap \neg\exists \mathsf{s} \quad \exists \mathsf{u}^- \sqsubseteq \exists \mathsf{r}^- \sqcap \neg\exists \mathsf{s}^- \quad \exists \mathsf{t} \sqsubseteq \exists \mathsf{s}^- \quad \mathsf{r} \sqsubseteq \mathsf{s}$$

and assign cost 1 to all t assertions, cost 2 to axiom $\mathsf{r} \sqsubseteq \mathsf{s}$, and infinite cost to other assertions and axioms. Every interpretation with cost ≤ 3 violates the assertion $\mathsf{t}(\mathsf{b}_0,\mathsf{c}_0)$, which is somewhat surprising as both involved individuals have ABox types that can otherwise be instantiated in a way that does not violate anything. However, b_0 and c_0 happen to be 'close', i.e. *at distance less than $k = 3$, to a_0. The rare type of a_0 can then impact b_0 and c_0 as seen above.*

To capture those individuals that may be affected by elements from the pre-core, we essentially explore the neighbourhood of the latter. For $\mathsf{a},\mathsf{b} \in \mathsf{Ind}(\mathcal{A})$, we write $\mathsf{a} \rightsquigarrow_1 \mathsf{b}$ if there exists a role $\mathsf{r} \in \mathsf{N}^\pm_{\mathsf{R}}$ and an assertion $\mathsf{r}(\mathsf{a},\mathsf{b})$ in $\mathcal{A}$ and $\exists^{>k}\mathsf{r} \notin \mathsf{tp}_{\mathcal{A}}(\mathsf{a})$. When exploring neighbours, the reason we exclude roles r such that $\exists^{>k}\mathsf{r} \in \mathsf{tp}_{\mathcal{A}}(\mathsf{a})$ comes from Lemma 7: it guarantees that element a satisfies $\exists\mathsf{r}$, so the potential violations on a do not impact the r-edges to b. We then denote $\mathsf{a} \rightsquigarrow_{i+1} \mathsf{c}$ if there exists b such that $\mathsf{a} \rightsquigarrow_i \mathsf{b}$ and $\mathsf{b} \rightsquigarrow_1 \mathsf{c}$. The core of $\mathcal{A}$, denoted $\mathsf{core}(\mathcal{A})$, is now defined as:

$$\mathsf{core}(\mathcal{A}) := \mathsf{pc}(\mathcal{A}) \cup \{\mathsf{b} \mid \mathsf{a} \rightsquigarrow_i \mathsf{b}, \mathsf{a} \in \mathsf{pc}(\mathcal{A}), i \leq k+1\}.$$

Notice that we stop the exploration of the neighbourhood of $\mathsf{pc}(\mathcal{A})$ at depth $k+1$. This is simply because the special behaviour of a pre-core element can only "cascade" to neighbours by enforcing a violation at each layer; thus, impacted elements cannot be further than $(k+1)$-away.

We can now state our key technical lemma. Note that Points 5_p and 5_c are used respectively to handle the possible and certain semantics. Recall that Theorem 7 established coNP-hardness for CQA_c^k, which is why Point 5_c only concerns the case where q is an IQ.

Lemma 8. *Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})_\omega$ be a WKB, q a BCQ, and k a fixed cost. If there exists an interpretation $\mathcal{I}$ with cost $\leq k$, then there exists an interpretation $\mathcal{J}$ such that:*

1. $\Delta^{\mathcal{J}} = \mathsf{Ind}(\mathcal{A}) \cup W$ *for some* $W \subseteq \{w_t \mid t \text{ is a 1-type}\}$;
2. $\mathcal{J}|_{\mathsf{pc}(\mathcal{A})} = \mathcal{I}|_{\mathsf{pc}(\mathcal{A})}$;
3. $\omega(\mathcal{J}) = \omega(\mathcal{J}|_{\mathsf{core}(\mathcal{A})\cup W})$;
4. $\omega(\mathcal{J}) \leq k$;

5_p. *If $\mathcal{I} \models q$, then $\mathcal{J}|_{\mathsf{core}(\mathcal{A})\cup W} \models q$ (and thus $\mathcal{J} \models q$);*

5_c. *If q is an IQ and $\mathcal{I} \not\models q$, then $\mathcal{J} \not\models q$.*

6.2 Construction of the FO-Rewriting

Consider a DL-Lite$_{\mathrm{bool}}^{\mathcal{H}}$ TBox $\mathcal{T}$, a weight function $\omega_{\mathcal{T}}$ for $\mathcal{T}$, fixed cost k, and BCQ q. We now proceed to the actual construction of the rewritten query q'. Notice that the pre-core always has size at most $P_0 := 2k \times 2^{|\mathsf{N_C}(\mathcal{T})|+4|\mathsf{N_R}(\mathcal{T})|} + |q|$ (at most $2k$ copies of each rare ABox type, plus the individuals from the IQ). For each individual a, there exist at most $2k|\mathsf{N_R}(\mathcal{T})|$ distinct individuals b such that $\mathsf{a} \rightsquigarrow_1 \mathsf{b}$. Therefore, the size of every core, regardless of the specific ABox, is bounded by $P_0 \times (2k|\mathsf{N_R}(\mathcal{T})|)^k$. We let $M_{\mathcal{T},q,k}$ be the above quantity plus the number of possible 1-types, that is:

$$M_{\mathcal{T},q,k} := P_0 \times (2k|\mathsf{N_R}(\mathcal{T})|)^k + 2^{|\mathsf{N_C}(\mathcal{T})|+2|\mathsf{N_R}(\mathcal{T})|}$$

The number $M_{\mathcal{T},q,k}$ gives an upper bound on the maximal size of the domain $\mathsf{core}(\mathcal{A}) \cup W$ involved in the construction of interpretation $\mathcal{J}$ in Lemma 8. Note that Points 3, 5_p and 5_c also guarantee that $\mathsf{core}(\mathcal{A}) \cup W$ contains all the relevant information regarding violations and query satisfaction.

The rewritten query q' considers each relevant interpretation $\mathcal{M}$, up to isomorphism, whose domain $\Delta^{\mathcal{M}}$ has size at most $M_{\mathcal{T},q,k}$ and tries to match the 'individual part' of $\mathcal{M}$ (that is, the $\mathsf{core}(\mathcal{A})$ part, as opposed to the W part) in the input $\mathcal{A}_{\omega_{\mathcal{A}}}$. The rewriting must also check that each remaining individual from $\mathcal{A}_{\omega_{\mathcal{A}}}$ can be interpreted in a manner that is compatible w.r.t. the considered $\mathcal{M}$, in the sense that it shall not introduce any violations (nor satisfy the query, in the case of the certain semantics). To achieve this, we specify along with $\mathcal{M}$ its intended individual part as a subdomain $\Gamma \subseteq \Delta^{\mathcal{M}}$ and the allowed ABox violations as a weighted ABox $\mathcal{V}_\nu$. We call such a triple $(\mathcal{M}, \Gamma, \mathcal{V}_\nu)$ a *strategy* for $(\mathcal{T}_{\omega_{\mathcal{T}}}, q, k)$ if $\mathsf{Ind}(\mathcal{V}) \subseteq \Gamma$ and $\omega_{\mathcal{T}}(\mathcal{M}) + \sum_{\alpha \in \mathcal{V}} \nu(\alpha) \leq k$.

We differentiate between p-strategies, used for CQA_p^k, and c-strategies, used for IQA_c^k. A *p-strategy* is a strategy $(\mathcal{M}, \Gamma, \mathcal{V}_\nu)$ that additionally satisfies $\mathcal{M} \models q$. An ABox type t is p-safe for a p-strategy $\sigma = (\mathcal{M}, \Gamma, \mathcal{V}_\nu)$ if there exists a 1-type t' such that:

1. for every $\mathsf{A} \in \mathsf{N_C}(\mathcal{T})$, if $\mathsf{A} \in t$, then $\mathsf{A} \in t'$;
2. for every $\mathsf{r} \in \mathsf{N_R^{\pm}}(\mathcal{T})$, if $\exists \mathsf{r} \in t$, then $\exists \mathsf{r} \in t'$;
3. t' does not violate any CIs from $\mathcal{T}$;
4. for every $\mathsf{r} \in \mathsf{N_R^{\pm}}(\mathcal{T})$, if $\exists \mathsf{r} \in t'$, then $\exists \mathsf{s} \in t'$ for every $\mathsf{s} \in \mathsf{N_R^{\pm}}(\mathcal{T})$ with $\mathsf{r} \sqsubseteq_{\mathcal{T}} \mathsf{s}$;
5. for every $\mathsf{r} \in \mathsf{N_R^{\pm}}(\mathcal{T})$, if $\exists \mathsf{r} \in t'$, then there exists $d \in \Delta^{\mathcal{M}}$ such that for every $\mathsf{s} \in \mathsf{N_R^{\pm}}(\mathcal{T})$ with $\mathsf{r} \sqsubseteq_{\mathcal{T}} \mathsf{s}$, we have $\exists \mathsf{s}^- \in \mathsf{tp}_{\mathcal{M}}(d)$.

Recall that we need only to define c-strategies for the case where q is an IQ. We may assume w.l.o.g. that q is a concept IQ[1]. A *c-strategy* for $(\mathcal{T}_{\omega_{\mathcal{T}}}, q, k)$ is a strategy $\sigma = (\mathcal{M}, \Gamma, \mathcal{V}_\nu)$ such that $\mathsf{Ind}(q) \subseteq \Gamma$ and $\mathcal{M} \not\models q$, and an ABox type t is c-safe for σ if there exists a 1-type t' that satisfies Conditions 1–5 of p-safe types, plus the following condition:

6. if $q = \exists y\, \mathsf{A}(y)$, then $\mathsf{A} \notin t'$.

The rewritten FO-query q' is now obtained as the disjunction of subqueries q_σ, where each σ is a p-strategy (resp. c-strategy) for $(\mathcal{T}_{\omega_{\mathcal{T}}}, q, k)$. Now, for a given $\sigma := (\mathcal{M}, \Gamma, \mathcal{V}_\nu)$, the subquery q_σ uses one existentially quantified variable v_d for each $d \in \Gamma$, and attempts to identify a subpart of the input weighted ABox $\mathcal{A}_{\omega_{\mathcal{A}}}$ that could be interpreted as $\mathcal{M}|_\Gamma$, up to the violations described in $\mathcal{V}_\nu$. For example, if $d \notin \mathsf{A}^{\mathcal{M}}$ and $\mathsf{A}(d) \in \mathcal{V}$, then q_σ contains the following subquery $\mathsf{A}(v_d) \rightarrow \mathsf{W}_{\mathsf{A}}^{\nu(\mathsf{A}(d))}(v_d)$, enforcing that variable v_d can only be mapped on an individual a such that $\mathsf{A}(\mathsf{a}) \in \mathcal{A}$ if $\omega_{\mathcal{A}}$ and ν agree on the cost of the assertion $\mathsf{A}(\mathsf{a})$. Using a universally quantified variable, q_σ also makes sure that all other individuals in $\mathcal{A}$ have an ABox type that is p-safe (resp. c-safe) w.r.t. σ. It is indeed clear, examining the very local requirements defining safe types that 'having a safe p- (or c-) type' can be verified using an appropriate FO-subquery.

In this manner, we can construct FO-queries q'_p and q'_c, respectively based on p- and c-strategies, such that the following properties hold, concluding the proof of Theorem 8:

Lemma 9. *Let $\mathcal{T}_{\omega_{\mathcal{T}}}$ be a weighted* DL-Lite$_{\mathrm{bool}}^{\mathcal{H}}$ *TBox, k an integer, and q a BCQ. For every weighted ABox $\mathcal{A}_{\omega_{\mathcal{A}}}$:*

$$(\mathcal{T}, \mathcal{A})_{\omega_{\mathcal{T}} \cup \omega_{\mathcal{A}}} \models_p^k q \text{ iff } \mathcal{I}_{\mathcal{A}_{\omega_{\mathcal{A}}}} \models q'_p.$$

Furthermore, if q is an IQ, then:

$$(\mathcal{T}, \mathcal{A})_{\omega_{\mathcal{T}} \cup \omega_{\mathcal{A}}} \not\models_c^k q \text{ iff } \mathcal{I}_{\mathcal{A}_{\omega_{\mathcal{A}}}} \models q'_c.$$

7 Conclusion

Our results significantly improve our understanding of the data complexity of query entailment under recently introduced cost-based semantics. In particular, we have proved a Δ_2^p upper bound for the optimal-cost certain and possible semantics, yielding tight complexity bounds for a wide range of lightweight and expressive DLs, up to $\mathcal{ALCHIO}$ and covering also prominent DL-Lite dialects. Moreover, we obtained surprising tractability results, showing that fixed-cost possible semantics (for CQs) and certain semantics (for IQs) in DL-Lite$_{\mathrm{bool}}^{\mathcal{H}}$ enjoy the same low AC^0 complexity as classical CQ answering in DL-Lite. We expect our upper bounds can be adapted to also handle negative role inclusions (to cover also DL-Lite$_{\mathcal{R}}$). For DLs with functionality or number restrictions, it does not suffice to work with finite interpretations, so wholly different methods are required. Developing a practical implementation of the FO-rewritings for the identified tractable cases is another interesting direction.

[1] IQs of the form $\exists y\, \mathsf{r}(\mathsf{a}, y)$ (resp. $\exists xy\, \mathsf{r}(x, y)$) can be handled by adding infinite-weight CIs $\exists \mathsf{r} \sqsubseteq \mathsf{B}, \mathsf{B} \sqsubseteq \exists \mathsf{r}$ for a fresh concept name B, and using the IQ $\mathsf{B}(\mathsf{a})$ (resp. $\exists x.\mathsf{B}(x)$) instead.

Acknowledgments

The authors acknowledge the financial support of the ANR AI Chair INTENDED (ANR-19-CHIA-0014) and the Federal Ministry of Research, Technology and Space of Germany and by Sächsische Staatsministerium für Wissenschaft, Kultur und Tourismus in the programme Center of Excellence for AI-research "Center for Scalable Data Analytics and Artificial Intelligence Dresden/Leipzig", project identification number: ScaDS.AI.

References

Baader, F.; Horrocks, I.; Lutz, C.; and Sattler, U. 2017. *An Introduction to Description Logic*. Cambridge University Press.

Bienvenu, M. 2020. A Short Survey on Inconsistency Handling in Ontology-Mediated Query Answering. *Künstliche Intell.*, 34(4): 443–451.

Bienvenu, M.; and Bourgaux, C. 2016. Inconsistency-Tolerant Querying of Description Logic Knowledge Bases. In *Reasoning Web Tutorial Lectures*.

Bienvenu, M.; Bourgaux, C.; and Jean, R. 2024a. Cost-Based Semantics for Querying Inconsistent Weighted Knowledge Bases. In *Proceedings of KR*.

Bienvenu, M.; Bourgaux, C.; and Jean, R. 2024b. Cost-Based Semantics for Querying Inconsistent Weighted Knowledge Bases (Long version with appendix). Http://arxiv.org/abs/2407.20754.

Bienvenu, M.; and Ortiz, M. 2015. Ontology-Mediated Query Answering with Data-Tractable Description Logics. In *Reasoning Web Tutorial Lectures*.

Bonatti, P. A.; Lutz, C.; and Wolter, F. 2009. The Complexity of Circumscription in DLs. *J. Artif. Intell. Res.*, 35: 717–773.

Bourgaux, C. 2016. *Inconsistency handling in ontology-mediated query answering. (Gestion des incohérences pour l'accès aux données en présence d'ontologies)*. Ph.D. thesis, University of Paris-Saclay, France.

Britz, K.; Casini, G.; Meyer, T.; Moodley, K.; Sattler, U.; and Varzinczak, I. 2021. Principles of KLM-style Defeasible Description Logics. *ACM Trans. Comput. Log.*, 22(1): 1:1–1:46.

Calvanese, D.; De Giacomo, G.; Lembo, D.; Lenzerini, M.; and Rosati, R. 2007. Tractable Reasoning and Efficient Query Answering in Description Logics: The DL-Lite Family. *Journal of Automated Reasoning (JAR)*.

Eiter, T.; Lukasiewicz, T.; and Predoiu, L. 2016. Generalized Consistent Query Answering under Existential Rules. In *Proceedings of KR*.

Giordano, L.; Gliozzi, V.; Olivetti, N.; and Pozzato, G. L. 2013. A non-monotonic Description Logic for reasoning about typicality. *Artif. Intell.*, 195: 165–202.

Krentel, M. W. 1988. The Complexity of Optimization Problems. *J. Comput. Syst. Sci.*, 36(3): 490–509.

Lembo, D.; Lenzerini, M.; Rosati, R.; Ruzzi, M.; and Savo, D. F. 2010. Inconsistency-Tolerant Semantics for Description Logics. In *Proceedings of RR*.

Lutz, C.; and Manière, Q. 2024. Adding Circumscription to Decidable Fragments of First-Order Logic: A Complexity Rollercoaster. In *Proceedings of KR*, 531–541.

Manière, Q. 2022. *Counting queries in ontology-based data access. (Requêtes de comptage pour l'accès aux données en présence d'ontologies)*. Ph.D. thesis, University of Bordeaux, France.

Ortiz, M.; Calvanese, D.; and Eiter, T. 2008. Data Complexity of Query Answering in Expressive Description Logics via Tableaux. *J. Autom. Reason.*, 41(1): 61–98.

Poggi, A.; Lembo, D.; Calvanese, D.; De Giacomo, G.; Lenzerini, M.; and Rosati, R. 2008. Linking Data to Ontologies. *Journal of Data Semantics*, 10: 133–173.

Xiao, G.; Calvanese, D.; Kontchakov, R.; Lembo, D.; Poggi, A.; Rosati, R.; and Zakharyaschev, M. 2018. Ontology-Based Data Access: A Survey. In *Proceedings of IJCAI*.

A Proofs for Section 3 (Cost-Based Semantics for Weighted KBs)

Lemma 1. *BCS^k for* DL-Lite$_{\text{core}}$ *(resp. $\mathcal{EL}_\bot$) reduces to IQA_p^k for* DL-Lite$_{\text{core}}$ *(resp. $\mathcal{EL}_\bot$). In particular, the same holds for BCS and IQA_p.*

Proof. Trivial: $\mathcal{K}_\omega$ is k-satisfiable iff $\mathcal{K}_\omega \models_p^k \mathsf{A}(\mathsf{a})$ where A is a fresh concept name and a is any individual name. □

Lemma 2. *IQA_p^k for* DL-Lite$_{\text{core}}$ *(resp. $\mathcal{EL}_\bot$) reduces to the complement of IQA_c^{k+1} for* DL-Lite$_{\text{core}}$ *(resp. $\mathcal{EL}_\bot$). In particular, IQA_p reduces to the complement of IQA_c.*

Proof. We first treat the case of DL-Lite$_{\text{core}}$. Let $(\mathcal{T}, \mathcal{A})_\omega$ be our DL-Lite$_{\text{core}}$ WKB, q our Boolean IQ and k our cost. Our reduction differs depending on the shape of q. We treat the three cases of (i) $q = \mathsf{A}(\mathsf{a})$, (ii) $q = \exists y\ \mathsf{A}(y)$ and (iii) $q = \mathsf{r}(\mathsf{a}, \mathsf{b})$, and argue that the other shapes are equivalent to an IQ of shape (i) or (ii) up to adding a few axioms in $\mathcal{T}$. Indeed, $q = \exists y\ \mathsf{r}(\mathsf{a}, y)$ is equivalent to $q' := \mathsf{B}(\mathsf{a})$ where B is a fresh concept name and we add the axiom $\mathsf{B} \equiv \exists \mathsf{r}$ to $\mathcal{T}$, with infinite weight. Similarly, $q = \exists y\ \mathsf{r}(y, \mathsf{a})$ is equivalent to $q' := \mathsf{B}(\mathsf{a})$ where B is a fresh concept name and we add the axiom $\mathsf{B} \equiv \exists \mathsf{r}^-$ to $\mathcal{T}$, with infinite weight. Finally, $q = \exists y_1\ \exists y_2\ \mathsf{r}(y_1, y_2)$ is equivalent to $q' := \exists y\ \mathsf{B}(y)$ where B is a fresh concept name and we add the axiom $\mathsf{B} \equiv \exists \mathsf{r}$ to $\mathcal{T}$, with infinite weight.

Now, for shape (i), that is $q = \mathsf{A}(\mathsf{a})$, the reduction proceeds as follows. We extend $\mathcal{T}$ into $\mathcal{T}'$ by adding axiom $\mathsf{A} \sqcap \bar{\mathsf{A}} \sqsubseteq \bot$ and $\mathcal{A}$ into $\mathcal{A}'$ by adding assertions $\mathsf{A}(\mathsf{a})$ and $\bar{\mathsf{A}}(\mathsf{a})$. The cost function ω is extended into ω' by assigning weight ∞ to fresh axiom $\mathsf{A} \sqcap \bar{\mathsf{A}} \sqsubseteq \bot$ and to assertion $\mathsf{A}(a)$, and weight 1 to $\bar{\mathsf{A}}(\mathsf{a})$.

Claim: $(\mathcal{T}, \mathcal{A})_\omega \models_p^k \mathsf{A}(\mathsf{a})$ iff $(\mathcal{T}', \mathcal{A}')_{\omega'} \not\models_c^{k+1} \bar{\mathsf{A}}(\mathsf{a})$.

($\Rightarrow$). Assume there exists an interpretation $\mathcal{I}$ that interprets the symbols in $(\mathcal{T}, \mathcal{A})$ in such a way that $\omega(\mathcal{I}) \leq k$ and $\mathcal{I} \models \mathsf{A}(\mathsf{a})$. Observe that if we define an interpretation $\mathcal{J}$ that is the same as $\mathcal{I}$ except that the concept $\bar{\mathsf{A}}$ is interpreted as $\bar{\mathsf{A}}^{\mathcal{J}} := \emptyset$, we obtain $\omega'(\mathcal{J}) \leq k + 1$. Moreover, since $\mathcal{J}$ does not satisfy $\bar{\mathsf{A}}(\mathsf{a})$, we get $(\mathcal{T}', \mathcal{A}')_{\omega'} \not\models_c^{k+1} \bar{\mathsf{A}}(\mathsf{a})$.
($\Leftarrow$). Assume there exists an interpretation $\mathcal{I}$ such that $\omega'(\mathcal{I}) \leq k + 1$ and $\mathcal{I} \not\models \bar{\mathsf{A}}(\mathsf{a})$. Both $\mathsf{A} \sqcap \bar{\mathsf{A}} \sqsubseteq \bot$ and $\mathsf{A}(\mathsf{a})$ have infinite weight w.r.t. ω', and thus are always satisfied in $\mathcal{I}$, as $\omega'(\mathcal{I})$ is finite. In particular $\mathcal{I} \models \mathsf{A}(a)$ and $\mathcal{I} \not\models \bar{\mathsf{A}}(\mathsf{a})$. It follows that $\omega'(\mathcal{I}) \geq \omega(\mathcal{I}) + \omega'(\bar{\mathsf{A}}(\mathsf{a}))$, thus $k + 1 - 1 \geq \omega(\mathcal{I})$, which yields $k \geq \omega(\mathcal{I})$. Therefore $(\mathcal{T}, \mathcal{A})_\omega \models_p^k \mathsf{A}(\mathsf{a})$.

We next consider the case of queries having shape (ii), where $q = \exists y\ \mathsf{A}(y)$. We extend $\mathcal{T}$ into $\mathcal{T}'$, $\mathcal{A}$ into $\mathcal{A}'$ and ω into ω' by adding the following axioms and assertion, with corresponding weights being indicated below:

$\mathsf{A}_0 \sqsubseteq \exists \mathsf{s}$	$\exists \mathsf{s}^- \sqsubseteq \mathsf{A}$	$\mathsf{A} \sqsubseteq \exists \mathsf{r}$	$\exists \mathsf{r}^- \sqsubseteq \bar{\mathsf{A}}$	$\mathsf{A}_0(\mathsf{a})$
∞	∞	∞	1	∞

where $\bar{\mathsf{A}}$, r, s and a are fresh.

Claim: $(\mathcal{T}, \mathcal{A})_\omega \models_p^k \exists y\ \mathsf{A}(y)$ iff $(\mathcal{T}', \mathcal{A}')_{\omega'} \not\models_c^{k+1} \exists y\ \bar{\mathsf{A}}(y)$.

($\Rightarrow$). Assume there exists an interpretation $\mathcal{I}$ which interprets the symbols in $(\mathcal{T}, \mathcal{A})_\omega$ in such a way that $\omega(\mathcal{I}) \leq k$ and $\mathsf{A}^{\mathcal{I}} \neq \emptyset$. We transform $\mathcal{I}$ into $\mathcal{J}$ by interpreting the fresh predicates $\bar{\mathsf{A}}$, r and s as follows:

$$\bar{\mathsf{A}}^{\mathcal{J}} := \emptyset \qquad \mathsf{r}^{\mathcal{J}} := \mathsf{A}^{\mathcal{I}} \times \{\mathsf{a}\} \qquad \mathsf{s}^{\mathcal{J}} := \{\mathsf{a}\} \times \mathsf{A}^{\mathcal{I}}$$

It is readily verified that $\omega'(\mathcal{I}) \leq k + 1$ (the extra cost coming from $\exists \mathsf{r}^- \sqsubseteq \bar{\mathsf{A}}$ being violated at a). Furthermore, as $\bar{\mathsf{A}}^{\mathcal{J}} := \emptyset$, we obtain $(\mathcal{T}', \mathcal{A}')_{\omega'} \not\models_c^{k+1} \exists y\ \bar{\mathsf{A}}(y)$.
($\Leftarrow$). Assume there exists an interpretation $\mathcal{I}$ such that $\omega'(\mathcal{I}) \leq k + 1$ and $\bar{\mathsf{A}}^{\mathcal{I}} = \emptyset$. By the three fresh axioms added to $\mathcal{T}'$ and the fresh assertion that have infinite cost, we can derive that $\mathsf{A}^{\mathcal{I}} \neq \emptyset$ and that $\exists \mathsf{r}^- \sqsubseteq \bar{\mathsf{A}}$ is violated at least once in $\mathcal{I}$. Observing that the cost of $\mathcal{I}$ w.r.t. the WKB $(\mathcal{T}, \mathcal{A})_\omega$ is at most k, we obtain $(\mathcal{T}, \mathcal{A})_\omega \models_p^k \exists y\ \mathsf{A}(y)$.

For shape (iii), that is $q = \mathsf{r}(\mathsf{a}, \mathsf{b})$, we extend $\mathcal{T}$ into $\mathcal{T}'$, $\mathcal{A}$ into $\mathcal{A}'$ and ω into ω' by adding the following axioms and assertion, with their corresponding weights being indicated below:

$\bar{\mathsf{A}} \sqcap \exists \mathsf{r} \sqsubseteq \bot$	$\mathsf{r}(\mathsf{a}, \mathsf{b})$	$\bar{\mathsf{A}}(\mathsf{a})$
∞	∞	1

We omit the proof of the following claim, which is essentially the same as the one used for shape (i):
Claim: $(\mathcal{T}, \mathcal{A})_\omega \models_p^k \mathsf{r}(\mathsf{a}, \mathsf{b})$ iff $(\mathcal{T}', \mathcal{A}')_{\omega'} \not\models_c^{k+1} \bar{\mathsf{A}}(\mathsf{a})$.

Let us now turn to the case of $\mathcal{EL}_\bot$. We can argue as before that it is sufficient to focus on the three same cases. Note that we need to change the argument for IQs with shape $q = \exists y\ \mathsf{r}(y, \mathsf{a})$, as we can no longer rely on inverse roles in the TBox. However, this lack of inverse roles in the ontology also means that the above q is actually equivalent to $q' = \exists y_1\ \exists y_2\ \mathsf{r}(y_1, y_2)$, which is turn equivalent to an IQ of shape (ii) as previously seen.

It is clear that we can reuse the same construction as before to handle queries of shapes (i) and (iii). For shape (ii), we simply replace the use of inverse roles by qualified existential restrictions. The added axioms in $\mathcal{T}$ become instead:

$$\begin{array}{ccc} \mathsf{A}_0 \sqsubseteq \exists \mathsf{s}.\mathsf{A} & \mathsf{A} \sqsubseteq \exists \mathsf{r}.\mathsf{A}_1 & \mathsf{A}_1 \sqsubseteq \bar{\mathsf{A}} \\ \infty & \infty & 1 \end{array}$$

□

Lemma 3. *For every $k \geq 0$, IQA_p^k for* DL-Lite$_{\text{core}}$ *(resp. $\mathcal{EL}_\bot$) reduces to IQA_p^{k+1} for* DL-Lite$_{\text{core}}$ *(resp. $\mathcal{EL}_\bot$) WKBs.*

Proof. Simply add an assertion $\mathsf{A}(\mathsf{a})$ with weight 1 and an axiom $\mathsf{A} \sqsubseteq \bot$ with weight ∞, where A and a are fresh. □

B Proofs for Section 4 (An Upper Bound for the General Case)

Theorem 1. *CQA_p (resp. CQA_c) for $\mathcal{ALCHIO}$ is in* NP *(resp. in* coNP*).*

Proof. An algorithm for CQA$_p$ (resp. CQA$_c$): (i) non-deterministically guesses an interpretation $\mathcal{I}$ whose size is bounded as in Lemma 4 (resp. Lemma 5), (ii) verifies whether $\mathcal{I} \models q$ (resp. $\mathcal{I} \not\models q$) and whether $\omega(\mathcal{I}) \leq k$; and, (iii) accepts iff both tests are successful. Note that both tests in step (ii) can be carried in deterministic polynomial time, so that the overall procedure yields an NP (resp. coNP) complexity upper bound. Soundness of such an algorithm is trivial, while its completeness is guaranteed by Lemma 4 (resp. Lemma 5). □

Theorem 2. *CQA_p^{opt} and CQA_c^{opt} for $\mathcal{ALCHIO}$ are in Δ_2^p.*

Proof. Notice that if the optimal cost is infinite, then the answer is trivially 'true' (resp. 'false') for CQA$_p$ (resp. CQA$_c$); and this case can easily be detected using an NP oracle (see *e.g.* (Ortiz, Calvanese, and Eiter 2008)) asking for the (standard) satisfiability of the sub-KB consisting of all assertions and axioms with infinite weight. Otherwise, the optimal cost is finite, and we observe that it is bounded from above by an exponential w.r.t. data complexity. Indeed, starting from $\mathcal{I}$ with finite cost k and q a tautology, we apply Lemma 4 and obtain an interpretation $\mathcal{J}$ whose size is bounded by a polynomial p in $|\mathcal{A}|$. Notice that $\mathcal{J}$ does not violate any assertion or axiom that has infinite cost since $\omega(\mathcal{J}) \leq k$. Let $W := \max_{\tau \in \mathcal{K}, \omega(\tau) < \infty}(\omega(\tau))$ be the maximal finite cost assigned by ω. Even if $\mathcal{J}$ violates all finite-cost assertions and axioms in every possible way, then $\omega(\mathcal{J}) \leq W(\left|\Delta^{\mathcal{J}}\right||\mathcal{A}| + \left|\Delta^{\mathcal{J}}\right|^2|\mathcal{T}|)$, which is at most an exponential w.r.t. data complexity (if the weights on $\mathcal{A}$ were encoded in unary, then it would only be a polynomial). Performing a binary search to identify the optimal cost thus only requires a polynomial number of calls to Theorem 1 (given again a tautology) seen as an NP-oracle. Once $optc(\mathcal{K}_\omega)$ is identified, we use Theorem 1 (this time with the query q of interest!) with input $optc(\mathcal{K}_\omega)$ as one last NP-oracle (resp. coNP-oracle) to decide CQA$_p^{opt}$ (resp. CQA$_c^{opt}$). □

Lemma 4. *Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})_\omega$ be an $\mathcal{ALCHIO}$ WKB, k an integer, and q a BCQ. If there exists an interpretation $\mathcal{I}$ such that $\omega(\mathcal{I}) \leq k$ and $\mathcal{I} \models q$, then there is an interpretation $\mathcal{J}$ such that $\omega(\mathcal{J}) \leq k$ and $\mathcal{J} \models q$, and whose domain $\Delta^{\mathcal{J}}$ has cardinality bounded polynomially in $|\mathcal{A}|$, with $|\mathcal{T}|$ and $|q|$ treated as constants, and independently from k.*

Proof. Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})_\omega$ be an $\mathcal{ALCHIO}$ WKB, k an integer, and q a BCQ. Assume we have an interpretation $\mathcal{I}$ with cost $\omega(\mathcal{I}) \leq k$ such that $\mathcal{I} \models q$. Consider the set of concepts occurring in the TBox $\mathcal{T}$ and close it under negation and sub-concepts. We denote $\mathcal{C}$ the obtained set of concepts, whose size is clearly bounded as $|\mathcal{C}| \leq 2|\mathcal{T}|^2$. The type $\mathsf{tp}_{\mathcal{I}}(d)$ of an element $d \in \Delta^{\mathcal{I}}$ is the unique subset of $\mathcal{C}$ such that, for every $C \in \mathsf{tp}_{\mathcal{I}}(d)$, we have $d \in C^{\mathcal{I}}$ iff $C \in \mathsf{tp}_{\mathcal{I}}(d)$. Note that this notion of type slightly differs from the one used in Section 6 and will only be used within the present proof. Let $W := \{w_{\mathsf{tp}_{\mathcal{I}}(d)} \mid d \in \Delta^{\mathcal{I}} \setminus \mathsf{Ind}\}$. We define the following mapping:

$$\begin{array}{rcl} \rho : \Delta^{\mathcal{I}} & \to & \mathsf{Ind}(\mathcal{K}) \cup W \\ d & \mapsto & \begin{cases} d & \text{if } d \in \mathsf{Ind}(\mathcal{K}) \\ w_{\mathsf{tp}_{\mathcal{I}}(d)} & \text{otherwise} \end{cases} \end{array}$$

where $\mathsf{Ind}(\mathcal{K})$ refers to the individuals occurring in the KB. We denote $\mathcal{J} := \rho(\mathcal{I})$ the image of interpretation $\mathcal{I}$ via ρ. Note that $\mathcal{J}$ has the desired size $|\mathsf{Ind}(\mathcal{K})| + |W| \leq |\mathcal{A}| + |\mathcal{T}| + 2^{2|\mathcal{T}|^2}$. Since q is a BCQ, it is preserved under homomorphisms, and thus we have $\mathcal{J} \models q$ from $\mathcal{I} \models q$. It remains to verify that $\omega(\mathcal{J}) \leq k$.

To this end, we first verify, as is standard with filtration, that:

$$(\star) \quad \text{for every } d \in \Delta^{\mathcal{I}}, \text{ we have } \mathsf{tp}_{\mathcal{J}}(\rho(d)) = \mathsf{tp}_{\mathcal{I}}(d)$$

We proceed by structural induction and show that for every $d \in \Delta^{\mathcal{I}}$, we have $\mathsf{C} \in \mathsf{tp}_{\mathcal{J}}(\rho(d))$ iff $\mathsf{C} \in \mathsf{tp}_{\mathcal{I}}(d)$.

- $\mathsf{C} \in \mathsf{N_C}$. Let $d \in \Delta^{\mathcal{I}}$. If $\mathsf{C} \in \mathsf{tp}_{\mathcal{I}}(d)$, then we have by definition of $\mathsf{C}^{\mathcal{J}}$ that $\mathsf{C} \in \mathsf{tp}_{\mathcal{J}}(\rho(d))$ as desired. Conversely, if $\mathsf{C} \in \mathsf{tp}_{\mathcal{J}}(\rho(d))$, then by definition of $\mathsf{C}^{\mathcal{J}}$ there exists $d' \in \Delta^{\mathcal{I}}$ such that $\mathsf{C} \in \mathsf{tp}_{\mathcal{I}}(d')$ and $\rho(d) = \rho(d')$. By definition of ρ, we get $\mathsf{tp}_{\mathcal{I}}(d) = \mathsf{tp}_{\mathcal{I}}(d')$ and thus $\mathsf{C} \in \mathsf{tp}_{\mathcal{I}}(d)$ as desired.
- $\mathsf{C} = \top$. By definition we have $\mathsf{C}^{\mathcal{I}} = \Delta^{\mathcal{I}}$ and $\mathsf{C}^{\mathcal{J}} = \Delta^{\mathcal{J}}$, thus $\mathsf{C} \in \mathsf{tp}_{\mathcal{J}}(\rho(d))$ iff $\mathsf{C} \in \mathsf{tp}_{\mathcal{I}}(d)$ trivially holds.

- $\mathsf{C} = \bot$. By definition we have $\mathsf{C}^{\mathcal{I}} = \emptyset$ and $\mathsf{C}^{\mathcal{J}} = \emptyset$, thus $\mathsf{C} \in \mathsf{tp}_{\mathcal{J}}(\rho(d))$ iff $\mathsf{C} \in \mathsf{tp}_{\mathcal{I}}(d)$ trivially holds.
- $\mathsf{C} = \{\mathsf{a}\}$. By definition we have $\mathsf{C}^{\mathcal{I}} = \mathsf{a}$ and $\mathsf{C}^{\mathcal{J}} = \mathsf{a}$. Notice that $\rho(\mathsf{a}) = \mathsf{a}$, thus $\mathsf{C} \in \mathsf{tp}_{\mathcal{J}}(\rho(d))$ if $\mathsf{C} \in \mathsf{tp}_{\mathcal{I}}(d)$. Conversely, by definition of ρ, the only pre-image of a by ρ is a, thus $\mathsf{C} \in \mathsf{tp}_{\mathcal{J}}(\rho(d))$ only if $\mathsf{C} \in \mathsf{tp}_{\mathcal{I}}(d)$.
- $\mathsf{C} = \neg\mathsf{B}$. Let $d \in \Delta^{\mathcal{I}}$. By the induction hypothesis, and using the fact that types are closed under taking subconcepts, we obtain that $\mathsf{B} \in \mathsf{tp}_{\mathcal{J}}(\rho(d))$ iff $\mathsf{B} \in \mathsf{tp}_{\mathcal{I}}(d)$. The semantics of concept negation gives $\mathsf{C}^{\mathcal{I}} = \Delta^{\mathcal{I}} \setminus \mathsf{B}^{\mathcal{I}}$ and $\mathsf{C}^{\mathcal{J}} = \Delta^{\mathcal{J}} \setminus \mathsf{B}^{\mathcal{J}}$, thus immediately $\mathsf{C} \in \mathsf{tp}_{\mathcal{J}}(\rho(d))$ iff $\mathsf{C} \in \mathsf{tp}_{\mathcal{I}}(d)$.
- $\mathsf{C} = \mathsf{B}_1 \sqcap \mathsf{B}_2$. The argument is similar to the one used for $\mathsf{C} = \neg\mathsf{B}$.
- $\mathsf{C} = \exists\mathsf{r}.\mathsf{B}$. Let $d \in \Delta^{\mathcal{I}}$. Assume first that $\mathsf{C} \in \mathsf{tp}_{\mathcal{I}}(d)$, that is, there exists $e \in \mathsf{B}^{\mathcal{I}}$ such that $(d, e) \in \mathsf{r}^{\mathcal{I}}$. Then, by definition of $\mathsf{r}^{\mathcal{J}}$, we have $(\rho(d), \rho(e)) \in \mathsf{r}^{\mathcal{J}}$. Furthermore, by the induction hypothesis and using the fact that types are closed under subconcepts, we obtain $\mathsf{B} \in \mathsf{tp}_{\mathcal{J}}(\rho(e))$, thus $\rho(d) \in (\exists\mathsf{r}.\mathsf{B})^{\mathcal{J}}$, that is, $\mathsf{C} \in \mathsf{tp}_{\mathcal{J}}(\rho(d))$ as desired. Conversely, assume that $\mathsf{C} \in \mathsf{tp}_{\mathcal{J}}(\rho(d))$, that is there exists $e \in \mathsf{B}^{\mathcal{J}}$ such that $(\rho(d), e) \in \mathsf{r}^{\mathcal{J}}$ and $e \in \mathsf{B}^{\mathcal{J}}$. By definition of $\mathsf{r}^{\mathcal{J}}$, there exists $(d', e') \in \mathsf{r}^{\mathcal{I}}$ such that $\rho(d') = \rho(d)$ and $\rho(e') = e$. By the induction hypothesis and using the fact that types are closed under subconcepts, we get $\mathsf{B} \in \mathsf{tp}_{\mathcal{I}}(e')$. Joint with $(d', e') \in \mathsf{r}^{\mathcal{I}}$, this gives $\mathsf{C} \in \mathsf{tp}_{\mathcal{I}}(d')$. Now using $\rho(d') = \rho(d)$, we obtain that $\mathsf{tp}_{\mathcal{I}}(d) = \mathsf{tp}_{\mathcal{I}}(d')$ and thus $\mathsf{C} \in \mathsf{tp}_{\mathcal{I}}(d)$ as desired.

We now come back to verifying $\omega(\mathcal{J}) \leq k$, and prove that $\omega(\mathcal{J}) \leq \omega(\mathcal{I})$, which suffices since the cost of $\mathcal{I}$ is assumed to be at most k. By $(\star)$, every concept assertion violated in $\mathcal{J}$ already is in $\mathcal{I}$. Similarly, every concept inclusion violated in $\mathcal{J}$ by an individual is already violated in $\mathcal{I}$ by the same individual. Still using $(\star)$, a violation of a concept inclusion in $\mathcal{J}$ on an element w_t is already violated by all non-individual elements of $\mathcal{I}$ whose type is t (note that by definition of W, there exists at least one). Since ρ is the identity on individuals, every role assertion violated in $\mathcal{J}$ already is in $\mathcal{I}$. It remains to treat the case of the violation at $(d, e) \in \mathsf{r}^{\mathcal{J}} \setminus \mathsf{s}^{\mathcal{J}}$ of a role inclusion $\mathsf{r} \sqsubseteq \mathsf{s}$. By definition of $\mathsf{r}^{\mathcal{J}}$, we have a pair $(d', e') \in \mathsf{r}^{\mathcal{I}}$ such that $\rho(d') = d$ and $\rho(e') = e$. By definition of $\mathsf{s}^{\mathcal{J}}$, there is no pair $(d'', e'') \in \mathsf{s}^{\mathcal{I}}$ such that $\rho(d'') = d$ and $\rho(e'') = e$ (as otherwise $(d, e) \in \mathsf{s}^{\mathcal{J}}$). Therefore the pair (d', e') already violates $\mathsf{r} \sqsubseteq \mathsf{s}$ in $\mathcal{I}$. Unfolding the definition of ρ, it is readily verified that the produced pair (d', e') is different for every different violation of $\mathsf{r} \sqsubseteq \mathsf{s}$. We proved that $\omega(\mathcal{J}) \leq \omega(\mathcal{I})$ as desired. □

Lemma 6. *Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ be an $\mathcal{ALCHIO}$ KB and q a BCQ. Let $\mathcal{V} \subseteq \mathcal{T}$ be a subset of $\mathcal{T}$ and $\mathcal{I}$ an interpretation such that $\mathcal{I} \not\models q$. There exists a polynomial p independent of $\mathcal{A}$, and an interpretation $\mathcal{J}$ satisfying the following:*

1. $\mathcal{J} \not\models q$; *3. $\forall \tau \in \mathcal{V}, vio_{\tau}(\mathcal{J}) \subseteq vio_{\tau}(\mathcal{I})$;*

2. $vio_{\mathcal{A}}(\mathcal{J}) = vio_{\mathcal{A}}(\mathcal{I})$; *4. $\left|\Delta^{\mathcal{J}}\right| \leq p(|\mathcal{A}| + |vio_{\mathcal{V}}(\mathcal{I})|)$.*

Before proceeding to the proof of Lemma 6, we briefly recall some standard terminology and notations that we will need. An interpretation $\mathcal{I}$ is a *model* of a KB $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ if it satisfies all inclusions in $\mathcal{T}$ and assertions in $\mathcal{A}$. We say that a query q is *entailed* from a KB $\mathcal{K}$, written $\mathcal{K} \models q$, if $\mathcal{I} \models q$ for every model $\mathcal{I}$ of $\mathcal{K}$.

A partial normal form. As the constructions in (Manière 2022) work under the usual semantics of models (not tolerating any violations) of KBs in normal form, our first step will be to replace the initial $\mathcal{ALCHIO}$ KB $\mathcal{K}$ with an $\mathcal{ALCHIO}$ KB in normal form that is able to keep track of violations of $\mathcal{K}$ while being an actual model. To do so, we rely on a minor adaptation of the normalization for $\mathcal{ALCO}$ WKBs proposed in (Bienvenu, Bourgaux, and Jean 2024b). Their construction removes all concept inclusions with finite weights and introduces new concept names into the TBox to be able to keep track of violations of those concept inclusions; they also prune the initial ABox to keep only those assertions that are satisfied in the interpretation $\mathcal{I}$. It is not hard to verify that inverse roles do not affect their translation into normal form. However, we cannot adopt their trick to keep track of violations of role inclusions as $\mathcal{ALCHIO}$ does not support expressive enough role inclusions, *e.g.* of the shape $\mathsf{r} \sqcap \neg\mathsf{s} \sqsubseteq \mathsf{t}$ where $\mathsf{r}, \mathsf{s}, \mathsf{t} \in \mathsf{N}_{\mathsf{R}}^{\pm}$. To circumvent this, we simply distinguish, in the original KB $\mathcal{K} := (\mathcal{T} \cup \mathcal{T}_{\mathcal{H}}, \mathcal{A})$, between the TBox $\mathcal{T}$ that only contains CIs of $\mathcal{ALCIO}$ concepts, and the remaining part $\mathcal{T}_{\mathcal{H}}$ of the TBox that contains all the RIs.

We now recall the normalization procedure proposed in (Bienvenu, Bourgaux, and Jean 2024b), phrased with inverse roles. Note that we drop their finite/infinite-cost-based distinction between parts of the WKB as we start from a KB without weights and simply work on violations.

We recall that a standard $\mathcal{ALCIO}$ KB is in normal form if all of its TBox axioms are in one of the following forms:

$$\top \sqsubseteq \mathsf{D} \qquad \mathsf{A} \sqsubseteq \mathsf{D} \qquad \{\mathsf{a}\} \sqsubseteq \mathsf{A} \qquad \mathsf{A}_1 \sqcap \mathsf{A}_2 \sqsubseteq \mathsf{A}$$
$$\exists\mathsf{r}.\mathsf{B} \sqsubseteq \mathsf{A} \qquad \mathsf{A} \sqsubseteq \exists\mathsf{r}.\mathsf{B} \qquad \neg\mathsf{B} \sqsubseteq \mathsf{A} \qquad \mathsf{A} \sqsubseteq \neg\mathsf{B}$$

where $\mathsf{A}, \mathsf{A}_1, \mathsf{A}_2, \mathsf{B} \in \mathsf{N}_{\mathsf{C}}$, $\mathsf{D} \in \mathsf{N}_{\mathsf{C}} \cup \{\{\mathsf{a}\} \mid \mathsf{a} \in \mathsf{N}_{\mathsf{I}}\}$, and $\mathsf{r} \in \mathsf{N}_{\mathsf{R}}^{\pm}$. It is well known that an arbitrary $\mathcal{ALCIO}$ TBox can be transformed (possibly through introduction of new concept names) into an $\mathcal{ALCIO}$ TBox in normal form. As the transformation may possibly introduce new concept names, the new TBox may not be equivalent to the original one, but it will be a conservative extension, which is sufficient for our purposes.

We proceed as follows to convert the $\mathcal{ALCIO}$ part $(\mathcal{T}, \mathcal{A})$ of the original KB $\mathcal{K} = (\mathcal{T} \cup \mathcal{T}_{\mathcal{H}}, \mathcal{A})$ into a KB in normal form:

1. For every $\tau = \mathsf{C} \sqsubseteq \mathsf{D} \in \mathcal{T}$, let A_{τ} be a fresh concept name (i.e. occurring neither in $\mathcal{K}$ nor q), and set

$$\mathcal{T}_{\tau} = \{\mathsf{A}_{\tau} \sqsubseteq \mathsf{C} \sqcap \neg\mathsf{D}, \mathsf{C} \sqcap \neg\mathsf{D} \sqsubseteq \mathsf{A}_{\tau}\}$$

2. Apply the normalization procedure to the TBox

$$\mathcal{T}^{vio} = \bigcup_{\tau \in \mathcal{T}} \mathcal{T}_\tau$$

and let $\mathcal{T}^\dagger$ denote the resulting TBox. We can assume w.l.o.g. that all concept names in $\mathsf{sig}(\mathcal{T}^\dagger) \setminus \mathsf{sig}(\mathcal{T}^{vio})$ occur neither in q nor $\mathcal{K}$.

We set $\mathcal{K}^\dagger = \langle \mathcal{T}^\dagger, \mathcal{A}^\dagger \rangle$, where

$$\mathcal{A}^\dagger = \{\alpha \in \mathcal{A} \mid \mathcal{I} \models \alpha\} \cup \{A_\top(\mathsf{a}) \mid \mathsf{a} \in \mathsf{Ind}(\mathcal{A})\}$$

with $\mathsf{A}_\top$ a fresh concept, not in $\mathcal{T}^\dagger$ nor q. Note that the assertions $\mathsf{A}_\top(\mathsf{a})$ simply serve to ensure that no individuals from $\mathcal{K}$ are lost during the translation (which will be convenient for the later constructions and proofs).

The following lemma summarizes the properties of $\mathcal{K}^\dagger$. Items (ii)-(iv) basically correspond to $\mathcal{T}^\dagger$ being a conservative extension of $\mathcal{T}^{vio}$, while (v) is a direct consequence of (iii) and the definition of $\mathcal{T}^{vio}$.

Lemma 10. *The KB $\mathcal{K}^\dagger$ satisfies the following:*

(i) $\mathsf{Ind}(\mathcal{K}) = \mathsf{Ind}(\mathcal{K}^\dagger)$
(ii) $\mathsf{sig}(\mathcal{T}) \subseteq \mathsf{sig}(\mathcal{T}^{vio}) \subseteq \mathsf{sig}(\mathcal{T}^\dagger)$*;*
(iii) every model of $\mathcal{T}^\dagger$ is a model of $\mathcal{T}^{vio}$;
(iv) for every model $\mathcal{I}_1$ of $\mathcal{T}^{vio}$, there exists a model $\mathcal{I}_2$ of $\mathcal{T}^\dagger$ with $\Delta^{\mathcal{I}_1} = \Delta^{\mathcal{I}_2}$ such that $\cdot^{\mathcal{I}_2}$ and $\cdot^{\mathcal{I}_1}$ coincide on all concept and role names except those in $\mathsf{sig}(\mathcal{T}^\dagger) \setminus \mathsf{sig}(\mathcal{T}^{vio})$*;*
(v) for every $\tau \in \mathcal{T}$ and every model $\mathcal{I}'$ of $\mathcal{K}^\dagger$, $vio_\tau(\mathcal{I}') = A_\tau^{\mathcal{I}'}$.

Using the preceding lemma, we can modify our original interpretation $\mathcal{I}$ to get a model of the KB $\mathcal{K}^\dagger$, that violates the RIs from $\mathcal{T}_\mathcal{H}$ exactly as $\mathcal{I}$ does:

Lemma 11. *There exists a model $\mathcal{I}^\dagger$ of $\mathcal{K}^\dagger$ with $\Delta^{\mathcal{I}^\dagger} = \Delta^{\mathcal{I}}$ such that $\cdot^{\mathcal{I}^\dagger}$ and $\cdot^{\mathcal{I}}$ coincide on* $\mathsf{sig}(\mathcal{K}) \cup \mathsf{sig}(q)$*. In particular, this means that:*

- $vio_\mathcal{A}(\mathcal{I}^\dagger) = vio_\mathcal{A}(\mathcal{I})$
- *for every $\tau \in \mathcal{T}$:* $vio_\tau(\mathcal{I}^\dagger) = vio_\tau(\mathcal{I})$
- *for every $\tau \in \mathcal{T}_\mathcal{H}$:* $vio_\tau(\mathcal{I}^\dagger) = vio_\tau(\mathcal{I})$
- $\mathcal{I}^\dagger \not\models q$

Proof. We first create an interpretation $\mathcal{I}^\diamond$ which is the same as $\mathcal{I}$ except for the concept names $\mathsf{A}_\tau \in \mathsf{sig}[\mathcal{T}^{vio}] \setminus \mathsf{sig}(\mathcal{T})$, for which we set $(\mathsf{A}_\tau)^{\mathcal{I}^\diamond} = (C \sqcap \neg D)^{\mathcal{I}}$, and the concept name $\mathsf{A}_\top$ which we interpret as $\mathsf{Ind}(\mathcal{A})$ (note that we may assume $\mathsf{Ind}(\mathcal{A}) \subseteq \Delta^{\mathcal{I}}$ due to the SNA on ABox individuals). This ensures that $\mathcal{I}^\diamond$ satisfies the axioms in $\mathcal{T}_\tau$ and the new assertions in $\mathcal{A}^\dagger$, and since we have not modified the interpretation of any other symbols, $\mathcal{I}^\diamond$ is a model of $\langle \mathcal{T}^{vio}, \mathcal{A}^\dagger \rangle$. Hence, by Lemma 10 (iv), there exists a model $\mathcal{I}^\dagger$ of $\mathcal{T}^\dagger$ which can be obtained from $\mathcal{I}^\diamond$ solely by changing the interpretation of concept names in $\mathsf{sig}(\mathcal{T}^\dagger) \setminus \mathsf{sig}(\mathcal{T}^{vio})$. In particular, $\mathcal{I}^\dagger$ is a model of $\mathcal{A}^\dagger$ and hence of the KB $\mathcal{K}^\dagger$. It follows from the construction of $\mathcal{I}^\dagger$ that $\Delta^{\mathcal{I}^\dagger} = \Delta^{\mathcal{I}}$ and that $\cdot^{\mathcal{I}^\dagger}$ and $\cdot^{\mathcal{I}}$ coincide on $\mathsf{sig}(\mathcal{K}) \cup \mathsf{sig}(q)$. The first, third, and fourth items follow immediately (for the third item, this is because $\cdot^{\mathcal{I}^\dagger}$ and $\cdot^{\mathcal{I}}$ coincide on $\mathsf{sig}(\mathcal{T}_\mathcal{H}) \subseteq \mathsf{sig}(\mathcal{K})$ and for the fourth item, recall that $\mathcal{I} \not\models q$). The second item holds due to $(\mathsf{A}_\tau)^{\mathcal{I}^\dagger} = (\mathsf{A}_\tau)^{\mathcal{I}^\diamond} = (\mathsf{C} \sqcap \neg \mathsf{D})^{\mathcal{I}}$. □

As the model $\mathcal{I}^\dagger$ of $\mathcal{K}^\dagger$ given in Lemma 5 satisfies the same key property as $\mathcal{I}$, i.e. not satisfying q, it can thus be used in place of $\mathcal{I}$ in the following constructions.

Construction of the interlacing With the normal form at hand, we now move towards the interlacing construction from (Manière 2022). The goal is to unravel the interpretation $\mathcal{I}^\dagger$ to obtain a more well-structured interpretation that retains the essential properties of $\mathcal{I}^\dagger$, i.e. it does not entail q and satisfies Points 2 and 3 in Lemma 6. The interlacing construction starts with the definition of the existential extraction, which is a tree-shaped domain. The original definition, which we adapt next, uses the alphabet $\Omega^\exists$ consisting of all $\mathsf{r}.\mathsf{A}$ such that $\exists \mathsf{r}.\mathsf{A}$ is the RHS of an axiom in the considered TBox. In the context of classical models of a KB, only keeping track of a role r of interest is sufficient as the considered interpretations being models, one can always infer that all roles s such that $\mathsf{r} \sqsubseteq_{\mathcal{T}_\mathcal{H}} \mathsf{s}$ will be satisfied whenever r is. To account for potentially violated RIs, we refine this alphabet by setting Ω to be the set of all $\mathsf{R}.\mathsf{A}$ such that $\mathsf{R} \subseteq \mathsf{N}_\mathsf{R}^\pm$ and $\mathsf{A} \in \mathsf{N}_\mathsf{C}$. Intuitively, the subset of roles R will keep track of which combination of roles is satisfied. If $\mathsf{R} \subseteq \mathsf{N}_\mathsf{R}^\pm$, we write $d \in (\exists \mathsf{R}.\mathsf{A})^{\mathcal{I}^\dagger}$ if there exists an element $e \in \Delta^{\mathcal{I}^\dagger}$, called an $\mathsf{R}.\mathsf{A}$-successor of d, such that:

- $e \in \mathsf{A}^{\mathcal{I}^\dagger}$; and
- $\mathsf{R} = \{\mathsf{r} \in \mathsf{N}_\mathsf{R}^\pm \mid (d, e) \in \mathsf{r}^{\mathcal{I}^\dagger}\}$.

Notice that this is in line with very expressive DLs that support more complex role constructors. Now, for every $\mathsf{R}.\mathsf{A} \in \Omega$, we assume chosen a function $\mathsf{succ}_{\mathsf{R}.\mathsf{A}}^{\mathcal{I}^\dagger}$ that maps every element $d \in \Delta^{\mathcal{I}^\dagger}$ that has a $\mathsf{R}.\mathsf{A}$-successor to one of its $\mathsf{R}.\mathsf{A}$-successors.

Definition 5 (Existential extraction)**.** *Build the following mapping inductively over the set* $\mathsf{Ind}(\mathcal{K}^\dagger)\cdot\Omega^*$*:*

$$\begin{array}{rcl} f : \mathsf{Ind}(\mathcal{K}^\dagger)\cdot\Omega^* & \to & \Delta^{\mathcal{I}^\dagger}\cup\{\uparrow\} \\ a & \mapsto & a \end{array}$$

$$w\cdot \mathsf{R.A} \quad \mapsto \quad \begin{cases} \uparrow & \text{if } f(w)=\uparrow \text{ or } f(w)\notin(\exists \mathsf{R.A})^{\mathcal{I}^\dagger} \\ \mathsf{succ}^{\mathcal{I}^\dagger}_{\mathsf{R.A}}(f(w)) & \text{otherwise} \end{cases}$$

where $\uparrow$ *is a fresh symbol witnessing the absence of a proper image for an element of* $\mathsf{Ind}(\mathcal{K}^\dagger)\cdot\Omega^*$*. The* existential extraction *of* $\mathcal{I}^\dagger$ *is* $\Delta^\circ := \{w \mid w\in\mathsf{Ind}(\mathcal{K}^\dagger)\cdot\Omega^*, f(w)\neq\uparrow\}$.

Apart from the extended notion of roles, which are now sets of roles, the preceding definition differs from the one in (Manière 2022) as it uses the set $\mathsf{Ind}(\mathcal{K}^\dagger)$ of individuals in the considered KB rather than the set of individuals in the considered ABox. This is because we work with $\mathcal{ALCHIO}$ whereas the original construction was formulated for $\mathcal{ALCHI}$. Fortunately, although nominals were not considered in (Manière 2022), they are properly handled by the constructions and do not require any notable modifications.

The original construction allows one to select a subdomain Δ^* of the starting interpretation $\mathcal{I}$ whose elements should not be duplicated when manipulating $\mathcal{I}$. This is where the parameter $\mathcal{V}$ in Lemma 6 comes into play: to guarantee Points 2 and 3, we should avoid duplicating violations of ABox assertions and of axioms from $\mathcal{V}$. We thus use the following definition of Δ^*:

$$\Delta^* := \mathsf{Ind}(\mathcal{K}) \cup \bigcup_{\tau\in\mathcal{V}} vio^d_\tau(\mathcal{I}^\dagger),$$

where $vio^d_\tau(\mathcal{I}^\dagger)$ refers to the *domain* where the violations of τ occurs: if τ is a CI, then $vio^d_\tau(\mathcal{I}^\dagger) := vio_\tau(\mathcal{I}^\dagger)$, otherwise τ is an RI and $vio^d_\tau(\mathcal{I}^\dagger) := \bigcup_{(e,e')\in vio_\tau(\mathcal{I}^\dagger)}\{e,e'\}$.

We can now introduce the function f^*, defined exactly as in (Manière 2022) but using the modified version of Δ^*:

$$\begin{array}{rcl} f^* : \quad \Delta^\circ & \to & \Delta^* \uplus (\Delta^\circ \setminus \Delta^*) \\ w & \mapsto & \begin{cases} f(w) & \text{if } f(w)\in\Delta^* \\ w & \text{otherwise} \end{cases} \end{array}$$

With these notions in hand, we are ready to recall the definition of the interlacing of $\mathcal{I}^\dagger$ w.r.t. the function f^*:

Definition 6 (f^*-interlacing)**.** *The* f^*-interlacing $\mathcal{I}^*$ *of* $\mathcal{I}^\dagger$ *is the interpretation whose domain is* $\Delta^{\mathcal{I}^*} := f^*(\Delta^\circ)$ *and which interprets concept and role names as follows:*

$$\begin{array}{rcl} \mathsf{A}^{\mathcal{I}^*} & := & \{f^*(u) \mid u\in\Delta^\circ, f(u)\in\mathsf{A}^{\mathcal{I}^\dagger}\} \\ \mathsf{r}^{\mathcal{I}^*} & := & \{(a,b) \mid a,b\in\mathsf{Ind}(\mathcal{K}^\dagger)\wedge(a,b)\in\mathsf{r}^{\mathcal{I}}\} \\ & & \cup\ \{(f^*(u), f^*(u\cdot\mathsf{R.B})) \mid u, u\cdot\mathsf{R.B}\in\Delta^\circ\wedge \mathsf{r}\in\mathsf{R}\} \\ & & \cup\ \{(f^*(u\cdot\mathsf{R.B}), f^*(u)) \mid u, u\cdot\mathsf{R.B}\in\Delta^\circ\wedge \mathsf{r}^-\in\mathsf{R}\}. \end{array}$$

The preceding definition is a slightly modified version of Definition 20 from (Manière 2022) as we read from the sets of roles R which role should hold between two elements, rather than systematically adding the super-roles of a specific role. We also make a slight modification – using $\mathsf{Ind}(\mathcal{K}^\dagger)$ rather than only individuals from the ABox – to make it compatible with nominals.

We have phrased the definition directly in terms of our desired function f^*, but other functions f' with domain Δ° can be used instead. Depending on which f' is used to define the interlacing, the resulting interpretation $\mathcal{I}'$ may or may not be a model of the considered KB. It was shown in (Manière 2022) that if f' is pseudo-injective, then the f'-interleaving is a model and moreover maps homomorphically into the starting interpretation. This property (stated in Lemma 12 below, phrased for our KB $\mathcal{K}^\dagger$) was proven for $\mathcal{ALCHI}$ KBs in normal form, but is easily shown to also hold for $\mathcal{ALCIO}$ KBs in normal form. In addition, it can be verified that our more refined interpretation of roles does not affect the homomorphism, but instead strengthens the connections between the interlacing and the starting interpretation $\mathcal{I}^\dagger$, as will be clarified in Lemma 14.

Definition 7. *A function* $f' : \Delta^\circ \to E$ *is* pseudo-injective *if: for all* $u, v\in\Delta^\circ$*, if* $f'(u) = f'(v)$*, then* $f(u) = f(v)$.

Lemma 12. *If* $f' : \Delta^\circ \to E$ *is pseudo-injective, then the* f'*-interlacing* $\mathcal{I}'$ *is a model of* $\mathcal{K}^\dagger$ *and the following mapping is a homomorphism from* $\mathcal{I}'$ *to* $\mathcal{I}^\dagger$*:*

$$\begin{array}{rcl} \sigma : \quad \Delta^{\mathcal{I}'} & \to & \Delta^{\mathcal{I}^\dagger} \\ f'(u) & \mapsto & f(u) \end{array}$$

As f' *is pseudo-injective,* σ *is well defined.*

It was proven in (Manière 2022) that f^* is pseudo-injective, and the same arguments apply to our modified f^*.

Lemma 13. *f^* is pseudo-injective.*

It follows from the preceding lemmas that $\mathcal{I}^*$ is a model of $\mathcal{K}^\dagger$ which maps homomorphically into $\mathcal{I}^\dagger$ via the mapping σ. In fact, due to the way f^* is defined, we can be more precise about the homomorphism σ:

Lemma 14. *The homomorphism σ from $\mathcal{I}^*$ to $\mathcal{I}^\dagger$ is such that $\sigma(\mathsf{a}) = \mathsf{a}$ for all $\mathsf{a} \in \mathsf{Ind}(\mathcal{K}^\dagger)$. Furthermore, for all role $\mathsf{r} \in \mathsf{N}_\mathsf{R}^\pm$ and $d, e \in \Delta^{\mathcal{I}^*}$, we have:*

$$(d, e) \in \mathsf{r}^{\mathcal{I}^*} \textit{ iff } (\sigma(d), \sigma(e)) \in \mathsf{r}^{\mathcal{I}^\dagger}$$

We are now ready to show that $\mathcal{I}^*$ retains the desired properties of $\mathcal{I}^\dagger$ (and thus of $\mathcal{I}$, by virtue of Lemma 11).

Lemma 15. *$\mathcal{I}^*$ satisfies Points 1, 2 and 3 of Lemma 6, that is:*

1. $\mathcal{I}^* \not\models q$;
2. $vio_\mathcal{A}(\mathcal{I}^*) = vio_\mathcal{A}(\mathcal{I}^\dagger)$;
3. $\forall \tau \in \mathcal{V}$, $vio_\tau(\mathcal{I}^*) \subseteq vio_\tau(\mathcal{I}^\dagger)$.

Proof. For Point 1, recall that Point 4 from Lemma 11 guarantees $\mathcal{I}^\dagger \not\models q$. The existence of the homomorphism σ from $\mathcal{I}^*$ to $\mathcal{I}^\dagger$ that is the identity on $\mathsf{Ind}(\mathcal{K}^\dagger)$ (see Lemmas 12 and 14), hence also on $\mathsf{Ind}(q)$, means that any match of q in $\mathcal{I}^*$ can be reproduced in $\mathcal{I}^\dagger$. Therefore, from $\mathcal{I}^\dagger \not\models q$, we immediately obtain $\mathcal{I}^* \not\models q$.

For Point 2, note that $vio_\mathcal{A}(\mathcal{I}^*) = vio_\mathcal{A}(\mathcal{I}^\dagger)$ is immediate from the definition of $\mathcal{I}^*$.

For Point 3, let $\tau \in \mathcal{V}$. We first treat the case of τ being a CI. Let $e \in vio_\tau(\mathcal{I}^*)$. Using Point (v) of Lemma 10 and $\mathcal{I}^*$ being a model of $\mathcal{K}^\dagger$, we obtain $e \in \mathsf{A}_\tau^{\mathcal{I}^*}$. By definition of $\mathsf{A}_\tau^{\mathcal{I}^*}$, there exists $d \in \Delta^\circ$ such that $f^*(d) = e$ and $f(d) \in \mathsf{A}_\tau^{\mathcal{I}^\dagger}$. Again using Point (v) of Lemma 10 but this time on $\mathcal{I}^\dagger$ being a model of $\mathcal{K}^\dagger$, we have $f(d) \in vio_\tau(\mathcal{I}^\dagger)$. As $vio_\tau(\mathcal{I}^\dagger) \subseteq \Delta^*$, we get $f^*(d) = f(d)$ by definition of f^*, thus $e = f^*(d) = f(d) \in vio_\tau(\mathcal{I}^\dagger)$ as desired. We turn to τ being a RI. Let $(d, e) \in vio_\tau(\mathcal{I}^*)$. By Lemma 14, applied on both roles occurring in the RI τ, we obtain $(\sigma(d), \sigma(e)) \in vio_\tau(\mathcal{I}^\dagger)$. As $vio_\tau(\mathcal{I}^\dagger) \subseteq \Delta^* \times \Delta^*$, we get $\sigma(d), \sigma(e) \in \Delta^*$, thus $\sigma(d) = d$ and $\sigma(e) = e$. We thus obtain $(d, e) \in vio_\tau(\mathcal{I}^\dagger)$ as desired. □

Quotient construction. It now remains to 'shrink' the interpretation $\mathcal{I}^*$ to obtain an interpretation with the same properties but having the required size. To do so, we can proceed exactly as in Chapter 3.4 of (Manière 2022), by first defining a suitable equivalence relation, then considering the quotient interpretation obtained by merging elements that belong to the same equivalence class.

We will now describe how the equivalence relation $\sim_n$ is defined, but without giving every detail as the definition is rather involved and we will not be modifying it except by using our own definition of Δ^*. The definition of $\sim_n$ involves the notion of the n-neighbourhood of an element d in $\mathcal{I}^*$ relative to Δ^*, denoted $\mathcal{N}_n^{\mathcal{I}^*,\Delta^*}(d)$, consisting of the elements in $\Delta^{\mathcal{I}^*}$ that can be reached by taking at most n 'steps' along role edges, starting at d and stopping whenever an element of Δ^* is reached. If $d \in \Delta^*$, then d itself is the only element in its n-neighbourhood (for any n). However, when $d \in \Delta^{\mathcal{I}^*} \setminus \Delta^*$, we know that $d = aw$ for some $a \in \mathsf{Ind}(\mathcal{K}^\dagger)$ and $w \in \Omega^*$. Using the tree-shaped structure of the domain Δ°, we can identify a unique 'root' prefix $r_{n,d}$ of $d = aw$ such that $f^*(r_{n,d}) \in \mathcal{N}_n^{\mathcal{I}^*,\Delta^*}(d)$ and for every $d' \in \mathcal{N}_n^{\mathcal{I}^*,\Delta^*}(d)$, there is a unique word $w_{n,d}^{d'} \in \Omega^*$ (note that Ω now refers to a full combination of roles) such that $d' = f^*(r_{n,d} \cdot w_{n,d}^{d'})$ and $|w_{n,d}^{d'}| \leq 2n$. With this uniform way of refer to the elements in neighbourhood of a considered element d, we can define a function $\chi_{n,d}$ whose output tells us for each word $w \in \Omega^*$ with $|w| \leq 2n$ whether there is an element in the neighbourhood whose word is w, and if so, whether that element belongs to Δ^*, and if not, which concept names it satisfies in $\mathcal{I}^*$. The equivalence relation $\sim_n$ then groups together those elements $d, e \in \Delta^{\mathcal{I}^*} \setminus \Delta^*$ which have the same associated word (i.e. $w_{n,d}^d = w_{n,e}^e$) and same associated function ($\chi_{n,d} = \chi_{n,e}$), plus an additional condition on the length of d and e (namely, $|d| = |e| \bmod 2|q| + 3$).

With the equivalence relation $\sim_n$ at hand, it now suffices to merge elements which are equivalent w.r.t. $\sim_{|q|+1}$. Formally, we consider the quotient interpretation $\mathcal{J} := \mathcal{I}^* / \sim_{|q|+1}$ whose domain is $\Delta^\mathcal{J} = \{[e]_{\sim_{|q|+1}} \mid e \in \Delta^{\mathcal{I}^*}\}$ and whose interpretation function $\cdot^\mathcal{J}$ is as follows:

$$a^\mathcal{J} = [a^{\mathcal{I}^*}]_{\sim_{|q|+1}} \qquad A^\mathcal{J} = \{[e]_{\sim_{|q|+1}} \mid e \in A^{\mathcal{I}^*}\} \qquad r^\mathcal{J} = \{([d]_{\sim_{|q|+1}}, [e]_{\sim_{|q|+1}}) \mid (d, e) \in r^{\mathcal{I}^*}\}$$

where $[e]_{\sim_{|q|+1}}$ is the equivalence class of e w.r.t. $\sim_{|q|+1}$.

It has been shown that $\mathcal{J}$ remains a model of the considered KB and does not contain any additional query matches. Given that the interpretation $\mathcal{I}^*$ we consider here is built in exactly the same way as the interleaving $\mathcal{I}'$ considered in (Manière 2022), except that we work with a larger domain of interest Δ^* (which includes the one used to build $\mathcal{I}'$), exactly the same arguments can be used to show Lemma 16 below. We point out that the presence of nominals in $\mathcal{K}^\dagger$ is not problematic as all of its individuals (whether present in the ABox or TBox) are included in Δ^* and are therefore left untouched by the construction.

Lemma 16. *The interpretation $\mathcal{J}$ is a model of $\mathcal{K}^\dagger$ such that $\mathcal{J} \not\models q$; in particular it satisfies Point 1 of Lemma 6.*

It remains to show that the quotient operation preserves the intended violations of assertions in $\mathcal{A}$ and of axioms from $\mathcal{V}$.

Lemma 17. *$\mathcal{J}$ satisfies Points 2 and 3 from Lemma 6.*

Proof. Notice that $\mathcal{J}$ and $\mathcal{I}^*$ coincide on Δ^*. Using Point 2 from Lemma 15 then yields the desired Point 2 for $\mathcal{J}$.

For Point 3, we distinguish between CIs from $\mathcal{V}$ and RIs from $\mathcal{V}$. For a CI $\tau \in \mathcal{V}$, we use $\mathcal{J}$ being a model of $\mathcal{K}^\dagger$ joint with Point (v) of Lemma 10, and the observation that $\mathcal{J}$ coincides with $\mathcal{I}^*$ on Δ^*, to obtain that $vio_\tau(\mathcal{J}) = vio_\tau(\mathcal{I}^\dagger)$. Now using Point 3 from Lemma 15 gives the desired result. For a RI $\mathsf{r} \sqsubseteq \mathsf{s}$, consider a pair $(d, e) \in \mathsf{r}^\mathcal{J}$ such that $(d, e) \notin \mathsf{s}^\mathcal{J}$. The definition of $\mathsf{r}^\mathcal{J}$ gives a pair $(d_0, e_0) \in \mathsf{r}^{\mathcal{I}^*}$ such that $[d_0]_{\sim_{|q|+1}} = d$ and $[e_0]_{\sim_{|q|+1}} = e$, and the definition of $\mathsf{s}^\mathcal{J}$ guarantees that $(d_0, e_0) \notin \mathsf{s}^{\mathcal{I}^*}$ (as otherwise (d, e) would be in $\mathsf{s}^\mathcal{J}$). It then follows from Lemma 15 that $d_0, e_0 \in \Delta^*$. Using the observation that $\mathcal{J}$ and $\mathcal{I}^*$ coincide on Δ^*, this yields $(d, e) \in vio_{\mathsf{r}\sqsubseteq\mathsf{s}}(\mathcal{I}^*)$, and we conclude using Point 3 of Lemma 15. □

To complete the proof of Lemma 6, we now examine the steps in the construction in order to place a bound on the size of the interpretation $\mathcal{J}$. By analyzing the number of equivalence classes, the following upper bound on $|\Delta^\mathcal{J}|$ was shown in (Manière 2022):

$$(2|q| + 3) \times |\mathcal{T}|^{|q|+2} \times (|\Delta^*| + 2^{\mathsf{sig}(\mathcal{T})} + 1)^{|\mathcal{T}|^{2|q|+2}}$$

Their analysis however relied on the size of the alphabet Ω being bounded by $\mathcal{T}$. In our case, Ω suffers from an exponential blow-up. To derive a naive upper bound in our setting, we simply replace each occurrence of $|\mathcal{T}|$ in the above by $2^{|\mathcal{T}|}$. Observe that the resulting upper bound remains a polynomial in Δ^*, hence a polynomial in $|\mathsf{Ind}(\mathcal{K})| + \left|vio^d_\mathcal{V}(\mathcal{I})\right|$. The latter term is upper bounded by $2|vio_\mathcal{V}(\mathcal{I})|$, and using $|\mathsf{Ind}(\mathcal{K})| \leq |\mathcal{T}| + |\mathcal{A}|$, one obtains the desired polynomial in $|\mathcal{A}| + |vio_\mathcal{V}(\mathcal{I})|$ whose coefficients are positive and independent of $\mathcal{A}$.

C Proofs for Section 5 (Lower Bounds)

Proofs for Section 5.1 (Lower Bounds in Extensions of $\mathcal{EL}_\bot$)

Theorem 3. *For every $k \geq 1$, BCS^k, IQA^k_p and CQA^k_p for $\mathcal{EL}_\bot$ are* NP*-hard.*

Proof. The reduction is from 3-SAT. Given a 3-CNF formula $\phi := \bigwedge_{i=1}^{\ell} \bigvee_{j=1}^{3} l_{i,j}$, where each $l_{i,j}$ is a literal over $v_1, \ldots, v_n$, we construct a WKB $(\mathcal{T}, \mathcal{A}_\phi)_\omega$. The ABox $\mathcal{A}_\phi$, encoding φ, contains the following assertions:

$$\begin{aligned}
\mathsf{False}(\mathsf{a}) \\
\mathsf{Bool}(\mathsf{a}) \\
\mathsf{Var}(v) &\text{ for each variable } v \\
\mathsf{clause}(\mathsf{a}, c_i) &\text{ for each } 1 \leq i \leq \ell \\
\mathsf{pos}_j(c_i, v) &\text{ for each } 1 \leq i \leq \ell \text{ and } 1 \leq j \leq 3 \text{ such that } l_{i,j} = v \\
\mathsf{neg}_j(c_i, v) &\text{ for each } 1 \leq i \leq \ell \text{ and } 1 \leq j \leq 3 \text{ such that } l_{i,j} = \neg v.
\end{aligned}$$

The TBox $\mathcal{T}$ has the following axioms:

$$\begin{array}{ccc}
\mathsf{Bool} \sqsubseteq \mathsf{True} & \mathsf{True} \sqcap \mathsf{False} \sqsubseteq \bot & \mathsf{Var} \sqsubseteq \exists \mathsf{val}.\mathsf{Bool} \\
\exists \mathsf{val}.\mathsf{True} \sqsubseteq \mathsf{True} & \exists \mathsf{val}.\mathsf{False} \sqsubseteq \mathsf{False} & \exists \mathsf{clause}.\mathsf{False} \sqsubseteq \mathsf{True} \\
\exists \mathsf{pos}_1.\mathsf{False} \sqcap \exists \mathsf{pos}_2.\mathsf{False} \sqcap \exists \mathsf{pos}_3.\mathsf{False} \sqsubseteq \mathsf{False} & & \exists \mathsf{pos}_1.\mathsf{False} \sqcap \exists \mathsf{pos}_2.\mathsf{False} \sqcap \exists \mathsf{neg}_3.\mathsf{True} \sqsubseteq \mathsf{False} \\
\exists \mathsf{pos}_1.\mathsf{False} \sqcap \exists \mathsf{neg}_2.\mathsf{True} \sqcap \exists \mathsf{pos}_3.\mathsf{False} \sqsubseteq \mathsf{False} & & \exists \mathsf{pos}_1.\mathsf{False} \sqcap \exists \mathsf{neg}_2.\mathsf{True} \sqcap \exists \mathsf{neg}_3.\mathsf{True} \sqsubseteq \mathsf{False} \\
\exists \mathsf{neg}_1.\mathsf{True} \sqcap \exists \mathsf{pos}_2.\mathsf{False} \sqcap \exists \mathsf{pos}_3.\mathsf{False} \sqsubseteq \mathsf{False} & & \exists \mathsf{neg}_1.\mathsf{True} \sqcap \exists \mathsf{pos}_2.\mathsf{False} \sqcap \exists \mathsf{neg}_3.\mathsf{True} \sqsubseteq \mathsf{False} \\
\exists \mathsf{neg}_1.\mathsf{True} \sqcap \exists \mathsf{neg}_2.\mathsf{True} \sqcap \exists \mathsf{pos}_3.\mathsf{False} \sqsubseteq \mathsf{False} & & \exists \mathsf{neg}_1.\mathsf{True} \sqcap \exists \mathsf{neg}_2.\mathsf{True} \sqcap \exists \mathsf{neg}_3.\mathsf{True} \sqsubseteq \mathsf{False}
\end{array}$$

The function ω assigns ∞ to all axioms and assertions, except for $\mathsf{Bool} \sqsubseteq \mathsf{True}$, which has weight 1.

Claim: ϕ is satisfiable iff $(\mathcal{T}, \mathcal{A}_\phi)_\omega$ is 1-satisfiable.

($\Leftarrow$). Assume ϕ is satisfiable, that is, there exists a valuation ν of $v_1, \ldots, v_n$ that makes ϕ true. We build an interpretation $\mathcal{I}$ of $(\mathcal{T}, \mathcal{A}_\phi)_\omega$ according to ν, with cost 1. The domain of $\mathcal{I}$ is $\Delta^\mathcal{I} := \mathsf{Ind}(\mathcal{A}_\phi) \cup \{b\}$, where b is a fresh domain element intended to represent the Boolean 'true', as opposed to individual a representing 'false'. The interpretation $\mathcal{I}$ interprets concept Var and roles clause, pos_j, neg_j as specified in the ABox. The remaining predicates are interpreted as follows:

$$\begin{aligned}
\mathsf{Bool}^\mathcal{I} &:= \{\mathsf{a}, b\} \\
\mathsf{False}^\mathcal{I} &:= \{\mathsf{a}\} \cup \{v_k \mid 1 \leq i \leq n, \nu(v_k) = 0\} \\
\mathsf{True}^\mathcal{I} &:= \{b\} \cup \{v_k \mid 1 \leq k \leq n, \nu(v_k) = 1\} \\
\mathsf{val}^\mathcal{I} &:= \{(v_k, \mathsf{a}) \mid 1 \leq k \leq n, \nu(v_k) = 0\} \cup \{(v_k, b) \mid 1 \leq k \leq n, \nu(v_k) = 1\}
\end{aligned}$$

It now suffices to verify that $\omega(\mathcal{I}) = 1$. All assertions from $\mathcal{A}_\phi$ are satisfied. Regarding axioms from $\mathcal{T}$, notice that a violates the concept inclusion $\mathsf{Bool} \sqsubseteq \mathsf{True}$, that has cost 1. The other axioms of $\mathcal{T}$ are all satisfied; in particular, since ν satisfies ϕ, it satisfies all its clauses, and thus none of the c_i's needs to satisfy the concept False.

($\Rightarrow$). Assume $(\mathcal{T}, \mathcal{A}_\phi)_\omega$ is 1-satisfiable. Consider an interpretation $\mathcal{I}$ with cost at most 1. Since assertion $\mathsf{False}(\mathsf{a})$ and axiom $\mathsf{True} \sqcap \mathsf{False}$ both have infinite cost, it is immediate that $\mathsf{a} \notin \mathsf{True}^\mathcal{I}$. From $\mathsf{Bool}(\mathsf{a}) \in \mathcal{A}_\phi$ having infinite cost and $\mathsf{Bool} \sqsubseteq \mathsf{True} \in \mathcal{T}$ having cost 1, the only violation in $\mathcal{I}$ is of the latter axiom, due to $\mathsf{a} \in \mathsf{Bool}^\mathcal{I} \setminus \mathsf{True}^\mathcal{I}$. In particular, all other axioms of $\mathcal{T}$ are satisfied by $\mathcal{I}$, and we also have $\mathsf{Bool}^\mathcal{I} \setminus \{\mathsf{a}\} \subseteq \mathsf{True}^\mathcal{I}$. Therefore $\mathcal{I}$ defines a valuation ν for variables $v_1, \ldots, v_n$: for each $1 \leq k \leq n$, we set $\nu(v_k) := 0$ if $(v_k, a) \in \mathsf{val}^\mathcal{I}$ (in which case $v_k \in (\exists \mathsf{val}.\mathsf{False})^\mathcal{I}$), and $\nu(v_k) = 1$ otherwise (in which case $v_k \in (\exists \mathsf{val}.\mathsf{True})^\mathcal{I}$). It is now readily verified that $\nu(\phi) = 1$ as otherwise we would have some $1 \leq i \leq \ell$ such that $c_i \in \mathsf{False}^\mathcal{I}$ (using one of the eight long axioms, depending on the shape of c_i), which would in turn yield $\mathsf{a} \in \mathsf{True}^\mathcal{I}$ (via axiom $\exists \mathsf{clause}.\mathsf{False} \sqsubseteq \mathsf{True}$), that is a contradiction. □

Theorem 4. *For every $k \geq 1$, IQA_c^k and CQA_c^k for $\mathcal{EL}$ are* coNP*-hard.*

Proof. We reduce from the problem of deciding whether a 3-DNF formula $\phi := \bigvee_{i=1}^{\ell} \bigwedge_{j=1}^{3} l_{i,j}$, where each $l_{i,j}$ is a variable from $v_1, \ldots, v_n$ or its negation, is a tautology. The IQ is $q := \mathsf{True}(\mathsf{a})$. The budget k is fixed to 1. The TBox has axioms:

$$\mathsf{Bool} \sqsubseteq \mathsf{True} \qquad \mathsf{Var} \sqsubseteq \exists \mathsf{val}.\mathsf{Bool}$$

$$\exists \mathsf{val}.\mathsf{True} \sqsubseteq \mathsf{True} \qquad \exists \mathsf{val}.\mathsf{False} \sqsubseteq \mathsf{False} \qquad \exists \mathsf{clause}.\mathsf{True} \sqsubseteq \mathsf{True}$$

$$\exists \mathsf{pos}_1.\mathsf{True} \sqcap \exists \mathsf{pos}_2.\mathsf{True} \sqcap \exists \mathsf{pos}_3.\mathsf{True} \sqsubseteq \mathsf{True} \qquad \exists \mathsf{pos}_1.\mathsf{True} \sqcap \exists \mathsf{pos}_2.\mathsf{True} \sqcap \exists \mathsf{neg}_3.\mathsf{False} \sqsubseteq \mathsf{True}$$

$$\exists \mathsf{pos}_1.\mathsf{True} \sqcap \exists \mathsf{neg}_2.\mathsf{False} \sqcap \exists \mathsf{pos}_3.\mathsf{True} \sqsubseteq \mathsf{True} \qquad \exists \mathsf{pos}_1.\mathsf{True} \sqcap \exists \mathsf{neg}_2.\mathsf{False} \sqcap \exists \mathsf{neg}_3.\mathsf{False} \sqsubseteq \mathsf{True}$$

$$\exists \mathsf{neg}_1.\mathsf{False} \sqcap \exists \mathsf{pos}_2.\mathsf{True} \sqcap \exists \mathsf{pos}_3.\mathsf{True} \sqsubseteq \mathsf{True} \qquad \exists \mathsf{neg}_1.\mathsf{False} \sqcap \exists \mathsf{pos}_2.\mathsf{True} \sqcap \exists \mathsf{neg}_3.\mathsf{False} \sqsubseteq \mathsf{True}$$

$$\exists \mathsf{neg}_1.\mathsf{False} \sqcap \exists \mathsf{neg}_2.\mathsf{False} \sqcap \exists \mathsf{pos}_3.\mathsf{True} \sqsubseteq \mathsf{True} \qquad \exists \mathsf{neg}_1.\mathsf{False} \sqcap \exists \mathsf{neg}_2.\mathsf{False} \sqcap \exists \mathsf{neg}_3.\mathsf{False} \sqsubseteq \mathsf{True}$$

The assertions are:

$$\begin{aligned}
\mathsf{False}(\mathsf{a}) \\
\mathsf{Bool}(\mathsf{a}) \\
\mathsf{Var}(v) &\text{ for each variable } v \\
\mathsf{clause}(\mathsf{a}, c_i) &\text{ for each } 1 \leq i \leq \ell \\
\mathsf{pos}_j(c_i, v) &\text{ for each } 1 \leq i \leq \ell \text{ and } 1 \leq j \leq 3 \text{ s.t. } l_{i,j} = v \\
\mathsf{neg}_j(c_i, v) &\text{ for each } 1 \leq i \leq \ell \text{ and } 1 \leq j \leq 3 \text{ s.t. } l_{i,j} = \neg v.
\end{aligned}$$

The function ω assigns ∞ to all axioms and assertions, except for $\mathsf{Bool} \sqsubseteq \mathsf{True}$, which has weight 1.

Claim: ϕ is a tautology iff $(\mathcal{T}, \mathcal{A}_\phi)_\omega \models_c^1 \mathsf{True}(a)$.

($\Leftarrow$). Assume ϕ is not a tautology, that is, there exists a valuation ν of $v_1, \ldots, v_n$ that makes ϕ false. We build an interpretation $\mathcal{I}$ of $(\mathcal{T}, \mathcal{A}_\phi)_\omega$ according to ν, with cost 1, and such that $\mathsf{a} \notin \mathsf{True}^\mathcal{I}$ (that is, $\mathcal{I} \not\models q$). The domain of $\mathcal{I}$ is $\Delta^\mathcal{I} := \mathsf{Ind}(\mathcal{A}_\phi) \cup \{b\}$, where b is a fresh domain element intended to represent the Boolean 'true', as opposed to individual a representing 'false'. The interpretation $\mathcal{I}$ interprets concept Var and roles clause, pos_j, neg_j as specified in the ABox. The remaining predicates are interpreted as follows:

$$\begin{aligned}
\mathsf{Bool}^\mathcal{I} &:= \{\mathsf{a}, b\} \\
\mathsf{False}^\mathcal{I} &:= \{\mathsf{a}\} \cup \{v_k \mid 1 \leq i \leq n, \nu(v_k) = 0\} \\
\mathsf{True}^\mathcal{I} &:= \{b\} \cup \{v_k \mid 1 \leq k \leq n, \nu(v_k) = 1\} \\
\mathsf{val}^\mathcal{I} &:= \{(v_k, \mathsf{a}) \mid 1 \leq k \leq n, \nu(v_k) = 0\} \cup \{(v_k, b) \mid 1 \leq k \leq n, \nu(v_k) = 1\}
\end{aligned}$$

It is immediate that $\mathsf{a} \notin \mathsf{True}^\mathcal{I}$, and it now suffices to verify that $\omega(\mathcal{I}) = 1$. All assertions from $\mathcal{A}_\phi$ are satisfied. Regarding axioms from $\mathcal{T}$, notice that a violates the concept inclusion $\mathsf{Bool} \sqsubseteq \mathsf{True}$, which has cost 1. The other axioms of $\mathcal{T}$ are all satisfied; in particular, since ν falsifies ϕ, it falsifies all its clauses, and thus none of the c_i's needs to satisfy the concept True.

($\Rightarrow$). Assume ϕ is a tautology. We need to prove that every interpretation with cost at most 1 satisfies q. Consider an interpretation $\mathcal{I}$ with cost at most 1. Note that $\mathsf{a} \in \mathsf{Bool}^\mathcal{I}$ as assertion $\mathsf{Bool}(\mathsf{a})$ has infinite cost. If $\mathsf{a} \in \mathsf{True}^\mathcal{I}$, then we are done. Otherwise, the cost of $\mathcal{I}$ is 1 due to the violation of the concept inclusion $\mathsf{Bool} \sqsubseteq \mathsf{True}$ at domain element a. In particular, all other axioms of $\mathcal{T}$ are satisfied by $\mathcal{I}$ and we also have $\mathsf{Bool}^\mathcal{I} \setminus \{\mathsf{a}\} \subseteq \mathsf{True}^\mathcal{I}$. From $\mathcal{I}$, we define a valuation of variables $v_1, \ldots, v_n$: : for each $1 \leq k \leq n$, we set $\nu(v_k) := 0$ if $(v_k, \mathsf{a}) \in \mathsf{val}^\mathcal{I}$ (in which case $v_k \in (\exists \mathsf{val}.\mathsf{False})^\mathcal{I}$), and $\nu(v_k) = 1$ otherwise (in which case $v_k \in (\exists \mathsf{val}.\mathsf{True})^\mathcal{I}$). By assumption, ϕ is a tautology, thus this valuation makes ϕ true. Therefore, one of its clauses c_i is satisfied by ν, which guarantees $c_i \in \mathsf{True}^\mathcal{I}$ due to one of the eight long axioms (depending on the shape of c_i). Axiom $\exists \mathsf{clause}.\mathsf{True} \sqsubseteq \mathsf{True}$ being satisfied then entails $\mathsf{a} \in \mathsf{True}^\mathcal{I}$, a contradiction. □

Proofs for Section 5.2 (Lower Bounds in the DL-Lite Family)

Theorem 5. *BCS, IQA$_p$ and CQA$_p$ for* DL-Lite$_{\text{core}}$ *are* NP*-hard. IQA$_c$ and CQA$_c$ for* DL-Lite$_{\text{core}}$ *WKBs are* coNP*-hard.*

Proof. We reduce from 3-COL, which is the problem of deciding whether a given graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is 3-colourable. All axioms of $\mathcal{T}$ are given infinite weight by ω and are as follows:

$$\exists s_i \sqcap \exists t_j \sqsubseteq \bot \quad \text{for } s,t \in \{\mathsf{r},\mathsf{g},\mathsf{b}\}, s \neq t, \text{ and } i,j \in \{1,2\}$$
$$\exists s_1^- \sqcap \exists s_2^- \sqsubseteq \bot \quad \text{for } s \in \{\mathsf{r},\mathsf{g},\mathsf{b}\}$$

We choose an orientation $\mathcal{E}'$ of $\mathcal{E}$, *i.e.* for each $\{u,v\} \in \mathcal{E}$, we add either (u,v) or (v,u) to $\mathcal{E}'$. For each $e = (u,v)$ in $\mathcal{E}'$ and each role name $s \in \{\mathsf{r},\mathsf{g},\mathsf{b}\}$, we add in the ABox $\mathcal{A}_\mathcal{G}$ the assertions $s_1(u,e)$ and $s_2(v,e)$. The assertions in $\mathcal{A}_\mathcal{G}$ are all given weight 1 by ω.

Claim: $\mathcal{G} \in$ 3-COL iff $(\mathcal{T}, \mathcal{A}_\mathcal{G})_\omega$ is $4|\mathcal{E}|$-satisfiable.

Notice that it is clear that every interpretation has cost at least $4|\mathcal{E}|$ as, for each edge $e \in \mathcal{E}'$, removing at least 4 of the 6 role assertions involving e is necessary to avoid the infinite cost of the TBox axioms.

($\Rightarrow$). Assume $\mathcal{G}$ is 3-colourable, that is, there exists a 3-colouring $\sigma : \mathcal{V} \to \{\mathsf{r},\mathsf{g},\mathsf{b}\}$. Consider the interpretation $\mathcal{I}$ that interprets, for $s \in \{\mathsf{r},\mathsf{g},\mathsf{b}\}$ and $i \in \{1,2\}$, the role name s_i as:

$$s_1^\mathcal{I} := \{(u,e) \mid e = (u,v) \in \mathcal{E}' \text{ and } \sigma(u) = s\}$$
$$s_2^\mathcal{I} := \{(v,e) \mid e = (u,v) \in \mathcal{E}' \text{ and } \sigma(v) = s\}$$

Note that the cost of ABox violations of $\mathcal{I}$ is exactly $4|\mathcal{E}|$. Furthermore, σ being a 3-colouring guarantees that $\mathcal{I}$ does not violate any TBox axiom. Therefore, $(\mathcal{T}, \mathcal{A}_\mathcal{G})_\omega$ is $4|\mathcal{E}|$-satisfiable.

($\Leftarrow$). Assume $(\mathcal{T}, \mathcal{A}_\mathcal{G})_\omega$ is $4|\mathcal{E}|$-satisfiable, that is, there exists an interpretation $\mathcal{I}$ with cost $\leq 4|\mathcal{E}|$, thus exactly $4|\mathcal{E}|$. We construct the mapping $\sigma : \mathcal{V} \to \{\mathsf{r},\mathsf{g},\mathsf{b}\}$ by setting $\sigma(v) = s$ iff $v \in (\exists s_1)^\mathcal{I} \cup (\exists s_2)^\mathcal{I}$. Notice that σ is a well-defined function as $\mathcal{I}$ has cost exactly $4|\mathcal{E}|$. It is then readily verified that σ is a 3-colouring of $\mathcal{G}$, hence $\mathcal{G} \in$ 3-COL. □

Theorem 6. *IQA$_p^{opt}$, IQA$_c^{opt}$, CQA$_p^{opt}$, and CQA$_c^{opt}$ for* DL-Lite$_{\text{core}}$ *are Δ_2^p-hard.*

Proof. The proof is by reduction from the following Δ_2^p-hard problem (Krentel 1988): given a satisfiable propositional 3-CNF formula $\varphi = c_1 \wedge \ldots \wedge c_m$ over variables $x_1, \ldots, x_n$ and given $k \in \{1, \ldots, n\}$, decide whether the lexicographically maximum truth assignment ν_{max} satisfying φ with respect to the ordering $(x_1, \ldots, x_n)$ is such that $\nu_{max}(x_k) = 1$. The reduction closely follows a similar reduction for querying DL-Lite KBs under preferred repair semantics (Bourgaux 2016, Proposition 6.2.4).

Given such a formula φ, we define a DL-Lite$_{\text{core}}$ WKB $\mathcal{K}_\omega = (\mathcal{T}, \mathcal{A})_\omega$ as follows:

$$\begin{aligned}
\mathcal{A} = &\{\mathsf{p}_\ell(c_j, x_i) \mid x_i \text{ is the } \ell\text{th literal of } c_j\} \cup \\
&\{\mathsf{n}_\ell(c_j, x_i) \mid \neg x_i \text{ is the } \ell\text{th literal of } c_j\} \cup \\
&\{\mathsf{T}(x_i) \mid 1 \leq i \leq n\} \\
\mathcal{T} = &\{\exists \mathsf{n}_\ell^- \sqsubseteq \neg \mathsf{T} \mid 1 \leq \ell \leq 3\} \cup \\
&\{\exists \mathsf{p}_\ell \sqsubseteq \neg \exists \mathsf{n}_{\ell'}, \exists \mathsf{p}_\ell^- \sqsubseteq \neg \exists \mathsf{n}_{\ell'}^- \mid 1 \leq \ell, \ell' \leq 3\} \cup \\
&\{\exists \mathsf{p}_\ell \sqsubseteq \neg \exists \mathsf{p}_{\ell'}, \exists \mathsf{n}_\ell \sqsubseteq \neg \exists \mathsf{n}_{\ell'} \mid 1 \leq \ell \neq \ell' \leq 3\}
\end{aligned}$$

We set $\omega(\tau) = \infty$ for every $\tau \in \mathcal{T}$. The weights of the ABox assertions will be defined through the following prioritization $\mathcal{P}$ which partitions $\mathcal{A}$ as follows:

- $L_1 = \mathcal{A} \setminus \{\mathsf{T}(x_i) \mid 1 \leq i \leq n\}$
- $L_p = \{\mathsf{T}(x_{p-1})\}$ for $1 < p \leq n+1$

following the method given in Lemma 6.2.5 in (Bourgaux 2016), which in our setting yields: $\omega(\alpha) = u^{n+1-p}$ for every $\alpha \in L_p$, where $u = 3m + 1$.

Let $\mathcal{I}$ be an interpretation with optimal cost. As the TBox assertions all have infinite weight, we know that $\mathcal{I} \models \mathcal{T}$. Furthermore, it follows from Proposition 3 of (BBJ 2024) that the ABox $\mathcal{A}_\mathcal{I} = \{\alpha \in \mathcal{A} \mid \mathcal{I} \models \alpha\}$ corresponding to the assertions satisfied in $\mathcal{I}$ is an $\leq_w$-repair, i.e. a subset of the ABox consistent with $\mathcal{T}$ and maximal among $\mathcal{T}$-consistent ABox subsets for the preorder defined by $\mathcal{A}_1 \leq_\omega \mathcal{A}_2$ if $\sum_{\alpha \in \mathcal{A}_1} \omega_\alpha \leq \sum_{\alpha \in \mathcal{A}_2} \omega_\alpha$. Moreover, as we have defined the weight of assertions following (Bourgaux 2016, Lemma 6.2.5) we know that $\mathcal{A}_\mathcal{I}$ is a $\leq_P$-repair of the KB $\langle \mathcal{T}, \mathcal{A} \rangle$ w.r.t. the prioritization $\mathcal{P}$. Formally, this means that there does not exist another subset $\mathcal{A}' \subseteq \mathcal{A}$ and priority level $1 \leq h < n+3$ such that (i) $\mathcal{A}'$ is $\mathcal{T}$-consistent, (ii) $|\mathcal{A}_\mathcal{I} \cap L_h| < |\mathcal{A}' \cap L_h|$, and (iii) $|\mathcal{A}_\mathcal{I} \cap L_h| = |\mathcal{A}' \cap L_h|$ for all $1 \leq g < h$.

It has been proven in (Bourgaux 2016, Proposition 6.2.4) that for every $\leq_P$-repair $\mathcal{A}^* \subseteq \mathcal{A}$ it is the case that $\mathsf{T}(x_i) \in \mathcal{A}^*$ iff the lexicographically maximum truth assignment ν_{max} satisfying φ fulfils $\nu_{max}(x_i) = 1$. In particular, this is the case for any ABox $\mathcal{A}_\mathcal{I}$ induced by an optimal-cost interpretation $\mathcal{I}$. It follows that for every optimal-cost interpretation $\mathcal{I}$, $\mathcal{I} \models \mathsf{T}(x_k)$ iff $\nu_{max}(x_k) = 1$. Hence $\mathcal{K}_\omega \models_c^{opt} \mathsf{T}(x_k)$ iff $\mathcal{K}_\omega \models_p^{opt} \mathsf{T}(x_k)$ iff $\nu_{max}(x_k) = 1$. □

Theorem 7. *For every $k \geq 1$, CQA_c^k for* DL-Lite$_{\text{core}}$ *is* coNP*-hard. This holds already for connected acyclic BCQs and without concept disjointness axioms.*

Proof. We reduce from deciding whether a 3-DNF formula $\phi := \bigvee_{i=1}^{\ell} \bigwedge_{j=1}^{3} l_{i,j}$, where each $l_{i,j}$ is a variable from $v_1, \ldots, v_n$ or its negation, is a tautology. We set $p_{i,j} := 0$ and $u_{i,j} := v_k$ if $l_{i,j} = \neg v_k$. Similarly, we set $p_{i,j} := 1$ and $u_{i,j} := v_k$ if $l_{i,j} = v_k$. The budget k is fixed to 1. All axioms and assertions have infinite weight except for the following, having weight 1:

$$\mathsf{Bool} \sqsubseteq \mathsf{True}.$$

The other axioms are:

$$\mathsf{Var} \sqsubseteq \exists \mathsf{val} \qquad\qquad \exists \mathsf{val}^- \sqsubseteq \mathsf{Bool}$$

The assertions are $\mathsf{False}(\mathsf{a})$, $\mathsf{True}(\mathsf{b})$, $\mathsf{Bool}(\mathsf{a})$, $\mathsf{Bool}(\mathsf{b})$, $\mathsf{val}(\mathsf{a},\mathsf{a})$, $\mathsf{val}(\mathsf{a},\mathsf{b})$, $\mathsf{val}(\mathsf{d},\mathsf{a})$ and:

$$\begin{aligned}
\mathsf{Var}(v_k) &\text{ for } 1 \leq k \leq n \\
\mathsf{clause}_s(\mathsf{a}_s, c_i) &\text{ for } 1 \leq i \leq \ell \text{ where } s := (p_{i,1}, p_{i,2}, p_{i,3}) \\
\mathsf{clause}_s(\mathsf{a}_{s'}, \mathsf{a}) &\text{ for } s, s' \in \{0,1\}^3 \text{ with } s \neq s' \\
\mathsf{clause}_s(\mathsf{d}, \mathsf{d}) &\text{ for } s \in \{0,1\}^3 \\
\mathsf{lit}_{j,p_{i,j}}(c_i, u_{i,j}) &\text{ for } 1 \leq i \leq \ell, 1 \leq j \leq 3 \\
\mathsf{lit}_{j,s}(\mathsf{a}, \mathsf{a}) &\text{ for } 1 \leq j \leq 3, s \in \{0,1\} \\
\mathsf{lit}_{j,s}(\mathsf{d}, \mathsf{d}) &\text{ for } 1 \leq j \leq 3, s \in \{0,1\}
\end{aligned}$$

The connected and acyclic BCQ q is the conjunction of 8 subqueries $q_s(y)$ for $s \in \{0,1\}^3$ that all share the variable y:

$$q := \exists y \bigwedge_{s \in \{0,1\}^3} q_s(y).$$

In turn, for each $s := (s_1, s_2, s_3) \in \{0,1\}^3$, the query $q_s(y)$ is defined as:

$$\begin{aligned}
q_s(y) = \; \exists y_{s,0}\, \exists y_{s,1}\, \exists y_{s,2}\, \exists y_{s,3}\, \exists y'_{s,1}\, \exists y'_{s,2}\, \exists y'_{s,3} \;\; & \mathsf{clause}_s(y, y_{s,0}) \wedge \bigwedge_{j=1}^{3} \left(\mathsf{lit}_{j,s_j}(y_{s,0}, y_{s,j}) \wedge \mathsf{val}(y_{s,j}, y'_{s,j}) \right) \\
& \wedge \bigwedge_{\substack{j \in \{1,2,3\} \\ s_j = 1}} \mathsf{True}(y'_{s,j}) \wedge \bigwedge_{\substack{j \in \{1,2,3\} \\ s_j = 0}} \mathsf{False}(y'_{s,j}).
\end{aligned}$$

Claim: ϕ is a tautology iff $(\mathcal{T}, \mathcal{A}_\phi)_\omega \models_c^1 q$.

($\Rightarrow$). Assume ϕ is a tautology. Consider an interpretation $\mathcal{I}$ with cost at most 1. First consider the case where $\mathsf{a} \in \mathsf{True}^{\mathcal{I}}$. In this case, we can trivially satisfy q by mapping all variables of the forms $y, y_{s,0}, y_{s,j}$ to d and all variables of the form $y'_{s,j}$ to a. Indeed, the ABox (which must be satisfied since the cost is at most 1 and ABox assertions have infinite weight) ensures that the role atoms are satisfied and that a belongs to False, and hence that a can be used to satisfy both True and False query atoms. Next consider the case where $\mathsf{a} \notin \mathsf{True}^{\mathcal{I}}$. In this case, the cost must be exactly 1, via a not being an instance of True despite $\mathsf{a} \in \mathsf{Bool}^{\mathcal{I}}$ (as the corresponding assertion has infinite weight) and $\mathsf{Bool} \sqsubseteq \mathsf{True}$ (the violation of which costs 1). Therefore, $(\exists \mathsf{val}^-)^{\mathcal{I}} \setminus \{\mathsf{a}\} \subseteq \mathsf{True}^{\mathcal{I}}$ as otherwise we would violate either $\exists \mathsf{val}^- \sqsubseteq \mathsf{Bool}$ or $\mathsf{Bool} \sqsubseteq \mathsf{True}$. We define the underlying valuation $\sigma : \{v_1, \ldots, v_n\} \to \{0,1\}$ by setting $\sigma(v_k) = 0$ if $(v_k, \mathsf{a}) \in \mathsf{val}^{\mathcal{I}}$ and $\sigma(v_k) = 1$ otherwise. Notice that, in the second case, that is $(v_k, \mathsf{a}) \notin \mathsf{val}^{\mathcal{I}}$, there exists at least an element $e \in \mathsf{True}^{\mathcal{I}}$ such that $(v_k, e) \in \mathsf{val}^{\mathcal{I}}$. For each such v_k, we choose one such e and denote it $f(v_k)$. Since ϕ is a tautology we have $\sigma(\phi) = 1$, thus there exists a $i \in \{1, \ldots, \ell\}$ such that $\sigma(c_i) = 1$. Let $s := (p_{i,1}, p_{i,2}, p_{i,3})$. We define a mapping h from variables of q to $\Delta^{\mathcal{I}}$ as follows:

$$\begin{aligned}
y &\mapsto a_s \\
y_{s,0} &\mapsto c_i \\
y_{s,j} &\mapsto u_{i,j} \\
y'_{s,j} &\mapsto \mathsf{a} && \text{if } \sigma(u_{i,j}) = 0 \\
y'_{s,j} &\mapsto f(u_{i,j}) && \text{if } \sigma(u_{i,j}) = 1 \\
y_{s',j} &\mapsto \mathsf{a} && \text{for } s' \neq s, j \in \{0,1,2,3\} \\
y'_{s',j} &\mapsto \mathsf{b} && \text{for } s' \neq s, j \in \{1,2,3\} \text{ s.t. } s'_j = 1 \\
y'_{s',j} &\mapsto \mathsf{a} && \text{for } s' \neq s, j \in \{1,2,3\} \text{ s.t. } s'_j = 0
\end{aligned}$$

It is readily verified that h is a homomorphism of q, thus $\mathcal{I} \models q$ as desired.

($\Leftarrow$). Assume ϕ is not a tautology. There exists a valuation ν of variables $v_1, \ldots, v_n$ that makes all clauses false. Let ν' be the "valuation" of variables $v_1, \ldots, v_n$ obtained by setting $\nu'(v) = \mathsf{a}$ if $\nu(v) = 0$ and $\nu'(v) = \mathsf{b}$ otherwise (in other words, just replace 0 by a and 1 by b in ν). We define an interpretation $\mathcal{I}$ corresponding to that valuation ν by interpreting all predicates as in the ABox, except for the role name val, that is interpreted as follows:

$$\mathsf{val}^{\mathcal{I}} := \{(\mathsf{a}, \mathsf{a}), (\mathsf{a}, \mathsf{b})\} \cup \{(v_k, \nu'(v_k)) \mid 1 \leq k \leq n\}.$$

It is immediate that the cost of $\mathcal{I}$ is 1, due to $\mathsf{a} \in \mathsf{Bool}^{\mathcal{I}} \setminus \mathsf{True}^{\mathcal{I}}$, violating the TBox axiom $\mathsf{Bool} \sqsubseteq \mathsf{True}$. It remains to verify that $\mathcal{I} \not\models q$. Assume by contradiction that there exists a homomorphism $h : q \to \mathcal{I}$. Due to the interpretation of the roles of the form clause_s, the variable y can only be mapped by h on one of the individuals a_s or on d. Note that the latter is excluded as $h(y) = \mathsf{d}$ would require $h(y'_{(1,1,1),1}) = \mathsf{a}$, to respect the interpretation of val, but $\mathsf{a} \notin \mathsf{True}^{\mathcal{I}}$. Therefore, $h(y) = a_s$ for some $s \in \{0,1\}^3$. Then the variable $y_{s,0}$ can only be mapped on some individual c_i for some $1 \leq i \leq \ell$ and, in turn, variables $y_{s,j}$ are mapped to individuals $u_{i,j}$. However, as ν makes all clauses false, and in particular c_i, it is readily verified that the images for variables $y'_{s,j}$ cannot adhere to the interpretation $\mathsf{val}^{\mathcal{I}}$ defined above. Therefore, $\mathcal{I} \not\models q$, thus $(\mathcal{T}, \mathcal{A}_\phi)_\omega \not\models^1_c q$. □

D Proofs for Section 6 (Positive Results in the DL-Lite Family)

Proofs for Section 6.1 (A (Very) Small Interpretation Property)

Lemma 7. *Consider an interpretation $\mathcal{I}$ whose cost is $\leq k$. For every individual a and role $\mathsf{r} \in \mathsf{N}_\mathsf{R}^\pm$, if $\exists^{>k}\mathsf{r} \in \mathsf{tp}_\mathcal{A}(\mathsf{a})$, then $\exists \mathsf{s} \in \mathsf{tp}_\mathcal{I}(\mathsf{a})$ for every role $\mathsf{s} \in \mathsf{N}_\mathsf{R}^\pm$ such that $\mathsf{r} \sqsubseteq_\mathcal{T} \mathsf{s}$.*

Proof. By contradiction, assume that there exists a role $\mathsf{s} \in \mathsf{N}_\mathsf{R}^\pm$ such that $\exists \mathsf{s} \notin \mathsf{tp}_\mathcal{I}(\mathsf{a})$. From $\exists^{>k}\mathsf{r} \in \mathsf{tp}_\mathcal{A}(\mathsf{a})$, we obtain $k+1$ distinct elements $\mathsf{b}_1, \ldots, \mathsf{b}_{k+1} \in \mathsf{N}_\mathsf{I}$ such that $\mathsf{r}(a, \mathsf{b}_1), \ldots, \mathsf{r}(a, \mathsf{b}_{k+1}) \in \mathcal{A}$ (we here treat the case of r not being an inverse role, the other case is similar). Note that for each such pair, we have $(a, \mathsf{b}_i) \notin \mathsf{s}^\mathcal{I}$ as $\exists \mathsf{s} \notin \mathsf{tp}_\mathcal{I}(\mathsf{a})$. We now prove that each pair (a, b_i) is involved in at least one binary violation, which will guarantee $k+1$ distinct violations, thus an overall cost for $\mathcal{I}$ that is $\geq k+1$, which is the desired contradiction. We distinguish based on $(a, \mathsf{b}_i) \in \mathsf{r}^\mathcal{I}$. If $(a, \mathsf{b}_i) \notin \mathsf{r}^\mathcal{I}$, then this is a clear violation of the role assertion $\mathsf{r}(a, \mathsf{b}_i)$. Else $(a, \mathsf{b}_i) \in \mathsf{r}^\mathcal{I}$, and from $(a, \mathsf{b}_i) \notin \mathsf{s}^\mathcal{I}$ we have $\mathsf{r} \neq \mathsf{s}$, thus at least one role inclusion used to witness $\mathsf{r} \sqsubseteq_\mathcal{T} \mathsf{s}$ is violated on (a, b_i). □

Lemma 8. *Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})_\omega$ be a WKB, q a BCQ, and k a fixed cost. If there exists an interpretation $\mathcal{I}$ with cost $\leq k$, then there exists an interpretation $\mathcal{J}$ such that:*

1. *$\Delta^\mathcal{J} = \mathsf{Ind}(\mathcal{A}) \cup W$ for some $W \subseteq \{w_t \mid t$ is a 1-type$\}$;*
2. *$\mathcal{J}|_{\mathsf{pc}(\mathcal{A})} = \mathcal{I}|_{\mathsf{pc}(\mathcal{A})}$;*
3. *$\omega(\mathcal{J}) = \omega(\mathcal{J}|_{\mathsf{core}(\mathcal{A}) \cup W})$;*
4. *$\omega(\mathcal{J}) \leq k$;*

5_p. *If $\mathcal{I} \models q$, then $\mathcal{J}|_{\mathsf{core}(\mathcal{A}) \cup W} \models q$ (and thus $\mathcal{J} \models q$);*

5_c. *If q is an IQ and $\mathcal{I} \not\models q$, then $\mathcal{J} \not\models q$.*

Figure 1 illustrates the construction underlying Lemma 8 and can be useful to follow its proof.

Proof. Let $\mathcal{K} = (\mathcal{T}, \mathcal{A})_\omega$ be a WKB, q a Boolean instance query, and k a fixed cost. Let $\mathcal{I}$ be an interpretation with cost $\leq k$. We define the set $\mathsf{crit}(\mathcal{I})$ of *critical elements* of $\mathcal{I}$ as the smallest set of elements from $\Delta^\mathcal{I}$ such that:

- $\mathsf{pc}(\mathcal{A}) \subseteq \mathsf{crit}(\mathcal{I})$; and
- if $d \in \mathsf{crit}(\mathcal{I})$, $d \rightsquigarrow_1 e$ and (d, e) is involved in a binary violation (*i.e.* either a role inclusion or a role assertion), then $e \in \mathsf{crit}(\mathcal{I})$.

Notice that, as the cost of $\mathcal{I}$ has cost $\leq k$ and that each inductive step in the second item above involves at least one violation, it is clear that $\mathsf{crit}(\mathcal{I}) \subseteq \mathsf{core}(\mathcal{A})$. We then set

$$W := \{w_t \mid t \in \mathsf{tp}_\mathcal{I}(\Delta^\mathcal{I} \setminus \mathsf{crit}(\mathcal{I}))\}$$

and define a mapping $\rho : \Delta^\mathcal{I} \to \mathsf{crit}(\mathcal{I}) \cup W$ as follows:

$$d \mapsto \begin{cases} d & \text{if } d \in \mathsf{crit}(\mathcal{I}) \\ w_{\mathsf{tp}_\mathcal{I}(d)} & \text{otherwise.} \end{cases}$$

For each individual $\mathsf{a} \in \mathsf{Ind}(\mathcal{A}) \setminus \mathsf{crit}(\mathcal{I})$, notice that $\mathsf{a} \notin \mathsf{pc}(\mathcal{A})$, thus the ABox type $t := \mathsf{tp}_\mathcal{A}(\mathsf{a})$ is not rare. As the cost of $\mathcal{I}$ is $\leq k$, there must exist some individual b such that $\mathsf{tp}_\mathcal{A}(\mathsf{b}) = t$ and b is not involved in any violation, that is all concept assertions and inclusions are satisfied by b and for all $d \in \Delta^\mathcal{I}$, all role assertions and inclusions are satisfied by (b, d) and at (d, b). For each such type t, we select one such 'perfect' individual p_t.

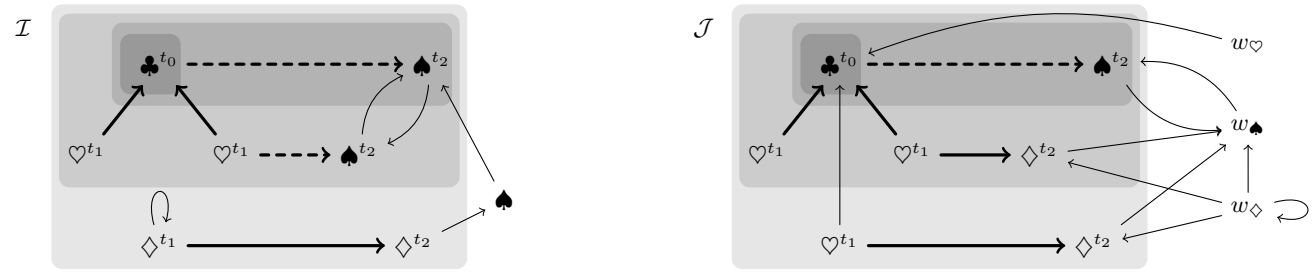

Figure 1: Two interpretations $\mathcal{I}$ and $\mathcal{J}$ of the same WKB. Symbols $\heartsuit, \diamondsuit$ indicate 1-types without unary violations; $\clubsuit, \spadesuit$ indicate 1-types with unary violations. An arrow indicates the role r; a bold arrow indicates a satisfied r-role assertion; a dashed bold arrow indicates the violation of an r-role assertion. Superscripts indicate the ABox type of an element; note that $\mathcal{I}$ features one anonymous element. We omit k and pretend t_0 is rare, while t_1 and t_2 are not (alternatively, one could set $k \geq 6$ and consider sufficiently many other individuals with ABox types t_1 and t_2). Shades of gray, from darker to lighter, highlight the pre-core, the critical elements, the core, and the individuals. $\mathcal{J}$ is obtained by applying Lemma 8 on $\mathcal{I}$, using the 'perfect' individual with 1-type $\heartsuit$ (resp. $\diamondsuit$) for ABox type t_1 (resp. t_2). Note that the *witness* $w_\spadesuit$ is given 1-type $\spadesuit$ and thus carries some unary violations.

We define $\mathcal{J}$ as the interpretation with domain $\mathsf{Ind}(\mathcal{A}) \cup W$ that interprets concept and role names as follows:

$$
\begin{aligned}
\mathsf{A}^{\mathcal{J}} &:= \rho(\mathsf{A}^{\mathcal{I}}) \cup \{\mathsf{a} \in \mathsf{Ind}(\mathcal{A}) \setminus \mathsf{crit}(\mathcal{I}) \mid p_{\mathsf{tp}_{\mathcal{A}}(\mathsf{a})} \in \mathsf{A}^{\mathcal{I}}\} \\
\mathsf{r}^{\mathcal{J}} &:= \rho(\mathsf{r}^{\mathcal{I}}) \\
&\quad \cup \left\{ (\mathsf{a},\mathsf{b}) \in \mathsf{Ind}(\mathcal{A})^2 \setminus \mathsf{crit}(\mathcal{I})^2 \;\middle|\; \begin{array}{l} \mathsf{s}(\mathsf{a},\mathsf{b}) \in \mathcal{A} \text{ for} \\ \text{some } \mathsf{s} \sqsubseteq_{\mathcal{T}} \mathsf{r} \end{array} \right\} \\
&\quad \cup \left\{ (\mathsf{b},\mathsf{a}) \in \mathsf{Ind}(\mathcal{A})^2 \setminus \mathsf{crit}(\mathcal{I})^2 \;\middle|\; \begin{array}{l} \mathsf{s}(\mathsf{a},\mathsf{b}) \in \mathcal{A} \text{ for} \\ \text{some } \mathsf{s} \sqsubseteq_{\mathcal{T}} \mathsf{r}^- \end{array} \right\} \\
&\quad \cup \{(\mathsf{a}, \rho(e)) \mid \mathsf{a} \in \mathsf{Ind}(\mathcal{A}) \setminus \mathsf{crit}(\mathcal{I}), (p_{\mathsf{tp}_{\mathcal{A}}(\mathsf{a})}, e) \in \mathsf{r}^{\mathcal{I}}\} \\
&\quad \cup \{(\rho(e), \mathsf{a}) \mid \mathsf{a} \in \mathsf{Ind}(\mathcal{A}) \setminus \mathsf{crit}(\mathcal{I}), (e, p_{\mathsf{tp}_{\mathcal{A}}(\mathsf{a})}) \in \mathsf{r}^{\mathcal{I}}\}
\end{aligned}
$$

Most of the work is to prove that the realized 1-types are indeed the expected ones, as shown by the next lemma:

Lemma 18. *For every* $\mathsf{a} \in \mathsf{crit}(\mathcal{I})$*, we have* $\mathsf{tp}_{\mathcal{J}}(\mathsf{a}) = \mathsf{tp}_{\mathcal{I}}(\mathsf{a})$*. For every* $w_t \in W$*, we have* $\mathsf{tp}_{\mathcal{J}}(w_t) = t$*. For every* $\mathsf{a} \in \mathsf{Ind}(\mathcal{A}) \setminus \mathsf{crit}(\mathcal{I})$*, we have* $\mathsf{tp}_{\mathcal{J}}(\mathsf{a}) = \mathsf{tp}_{\mathcal{I}}(p_{\mathsf{tp}_{\mathcal{A}}(\mathsf{a})})$.

Proof. Consider $\mathsf{a} \in \mathsf{crit}(\mathcal{I})$. Since, for every concept name A and role name r, we have $\rho(\mathsf{A}^{\mathcal{I}}) \subseteq \mathsf{A}^{\mathcal{J}}$ and $\rho(\mathsf{r}^{\mathcal{I}}) \subseteq \mathsf{r}^{\mathcal{J}}$, and ρ is the identity on $\mathsf{crit}(\mathcal{I})$, it is clear that $\mathsf{tp}_{\mathcal{I}}(\mathsf{a}) \subseteq \mathsf{tp}_{\mathcal{J}}(\mathsf{a})$. If $\mathsf{A} \in \mathsf{tp}_{\mathcal{J}}(\mathsf{a})$, it is also immediate that $\mathsf{A} \in \mathsf{tp}_{\mathcal{I}}(\mathsf{a})$. It remains to treat the case of $\exists \mathsf{r} \in \mathsf{tp}_{\mathcal{J}}(\mathsf{a})$. Assume there exists $(\mathsf{a}, d) \in \mathsf{r}^{\mathcal{J}}$. We distinguish five cases based on the definition of $\mathsf{r}^{\mathcal{J}}$. Case 4 is not applicable as $\mathsf{a} \in \mathsf{crit}(\mathcal{I})$. In both Cases 1 and 5, we directly obtain a witness for $\mathsf{a} \in (\exists \mathsf{r})^{\mathcal{I}}$, ensuring $\exists \mathsf{r} \in \mathsf{tp}_{\mathcal{I}}(\mathsf{a})$. We now turn to Case 2, that is $\mathsf{s}(\mathsf{a}, d) \in \mathcal{A}$ for some $\mathsf{s} \sqsubseteq_{\mathcal{T}} \mathsf{r}$. Note that $d \notin \mathsf{crit}(\mathcal{I})$, as otherwise we would have $(\mathsf{a}, d) \in \mathsf{crit}(\mathcal{I})^2$ and Case 2 would not be applicable. If $\mathsf{a} \rightsquigarrow_1 d$, then from $d \notin \mathsf{crit}(\mathcal{I})$ we obtain that there are no binary violations on (a, d), thus $(\mathsf{a}, d) \in \mathsf{s}^{\mathcal{I}} \subseteq \mathcal{I}$ and we are done. Otherwise $\mathsf{a} \rightsquigarrow_1 d$ is false, and in particular $\exists^{>k} \mathsf{s} \in \mathsf{tp}_{\mathcal{A}}(\mathsf{a})$. Therefore, $\mathsf{a} \in (\exists \mathsf{r})^{\mathcal{I}}$ by Lemma 7. Case 3 is treated as Case 2.

We now turn to the case of $w_t \in W$. Since, for every concept name A and role name r, we have $\rho(\mathsf{A}^{\mathcal{I}}) \subseteq \mathsf{A}^{\mathcal{J}}$ and $\rho(\mathsf{r}^{\mathcal{I}}) \subseteq \mathsf{r}^{\mathcal{J}}$, and that there exists a least one element $d \in \Delta^{\mathcal{I}} \setminus \mathsf{crit}(\mathcal{I})$ such that $\mathsf{tp}_{\mathcal{I}}(d) = t$, it is clear that $t \subseteq \mathsf{tp}_{\mathcal{J}}(w_t)$. If $\mathsf{A} \in \mathsf{tp}_{\mathcal{J}}(w_t)$, it is also immediate that $\mathsf{A} \in t$. It remains to treat the case of $\exists \mathsf{r} \in \mathsf{tp}_{\mathcal{J}}(w_t)$. Assume there exists $(w_t, d) \in \mathsf{r}^{\mathcal{J}}$. We distinguish 5 cases based on the definition of $\mathsf{r}^{\mathcal{J}}$. Cases 2, 3 and 4 are not applicable. Case 1 immediately provides a witness $d \in \Delta^{\mathcal{I}}$ with $\mathsf{tp}_{\mathcal{I}}(d) = t$ and such that $d \in (\exists \mathsf{r})^{\mathcal{I}}$. We turn to Case 5, that is $d \in \mathsf{Ind}(\mathcal{A}) \setminus \mathsf{crit}(\mathcal{I})$, and there exists $(e, p_{\mathsf{tp}_{\mathcal{A}}(d)}) \in \mathsf{r}^{\mathcal{I}}$ with $\rho(e) = w_t$. In particular, $\mathsf{tp}_{\mathcal{I}}(e) = t$, thus $e \in (\exists \mathsf{r})^{\mathcal{I}}$, as witnesses by $p_{\mathsf{tp}_{\mathcal{A}}(d)}$, proves $\exists \mathsf{r} \in t$.

The case of $\mathsf{a} \in \mathsf{Ind}(\mathcal{A}) \setminus \mathsf{crit}(\mathcal{I})$ is trivial, it suffices to recall that $\mathsf{tp}_{\mathcal{A}}(\mathsf{a}) = \mathsf{tp}_{\mathcal{A}}(p_{\mathsf{tp}_{\mathcal{A}}(\mathsf{a})})$ and that $p_{\mathsf{tp}_{\mathcal{A}}(\mathsf{a})}$ is not involved in any violations. □

Points 1 and 2 are trivially satisfied, by construction of $\mathcal{J}$. For Point 3, we prove a slightly stronger statement, namely that the cost of $\mathcal{J}$ equals the cost of $\mathcal{J}|_{\mathsf{crit}(\mathcal{I}) \cup W}$. Since $\mathsf{crit}(\mathcal{I}) \subseteq \mathsf{core}(\mathcal{A})$, Point 3 then follows. To prove the stronger statement, we need to guarantee that no element from $\mathsf{Ind}(\mathcal{A}) \setminus \mathsf{crit}(\mathcal{I})$ is involved in any violations. By virtue of the interpretation of such elements being dictated by the interpretation of the chosen elements p_t, which are not involved in any violations in $\mathcal{I}$, the claim

will follow. Let $\mathsf{a} \in \mathsf{Ind}(\mathcal{A}) \setminus \mathsf{crit}(\mathcal{I})$. For a concept assertion $\mathsf{A}(\mathsf{a}) \in \mathcal{A}$, recall that a and $p_{\mathsf{tp}_\mathcal{A}(\mathsf{a})}$ have the same ABox type, thus $\mathsf{A}(p_{\mathsf{tp}_\mathcal{A}(\mathsf{a})}) \in \mathcal{A}$. By definition of $p_{\mathsf{tp}_\mathcal{A}(\mathsf{a})}$, we have $p_{\mathsf{tp}_\mathcal{A}(\mathsf{a})} \in \mathsf{A}^\mathcal{I}$, thus, by definition of $\mathsf{A}^\mathcal{J}$, we obtain $\mathsf{a} \in \mathsf{A}^\mathcal{J}$. For a concept inclusion, it simply follows from Lemma 18 and the definition of $p_{\mathsf{tp}_\mathcal{A}(\mathsf{a})}$ not being involved in any violation. For a role assertion $\mathsf{r}(\mathsf{a},\mathsf{b}) \in \mathcal{A}$, note that we can directly use the definition of $\mathsf{r}^\mathcal{J}$ to obtain $(\mathsf{a},\mathsf{b}) \in \mathsf{r}^\mathcal{J}$. For a role inclusion $\mathsf{r} \sqsubseteq \mathsf{s}$, we treat the case of $(\mathsf{a},d) \in \mathsf{r}^\mathcal{J} \setminus \mathsf{s}^\mathcal{J}$, the case of $(d,\mathsf{a}) \in \mathsf{r}^\mathcal{J} \setminus \mathsf{s}^\mathcal{J}$ being symmetric. By definition of $\mathsf{r}^\mathcal{J}$, we have two options: either $d \in \mathsf{Ind}(\mathcal{A})$ with $\mathsf{t}(\mathsf{a},d) \in \mathcal{A}$ for some $\mathsf{t} \sqsubseteq_\mathcal{T} \mathsf{r}$, or $d = \rho(e)$ and $(p_{\mathsf{tp}_\mathcal{A}(\mathsf{a})}, e) \in \mathsf{r}^\mathcal{I}$. We start with $d \in \mathsf{Ind}(\mathcal{A})$ with $\mathsf{t}(\mathsf{a},d) \in \mathcal{A}$ for some $\mathsf{t} \sqsubseteq_\mathcal{T} \mathsf{r}$. In that case, we have $\mathsf{t} \sqsubseteq_\mathcal{T} \mathsf{s}$ since $\mathsf{t} \sqsubseteq_\mathcal{T} \mathsf{r}$ and $\mathsf{r} \sqsubseteq \mathsf{s} \in \mathcal{T}$, thus the definition of $\mathsf{s}^\mathcal{J}$ directly gives $(\mathsf{a},d) \in \mathsf{s}^\mathcal{J}$, which contradicts our assumption that $(\mathsf{a},d) \in \mathsf{r}^\mathcal{J} \setminus \mathsf{s}^\mathcal{J}$. It remains to turn to the case of $d = \rho(e)$ and $(p_{\mathsf{tp}_\mathcal{A}(\mathsf{a})}, e) \in \mathsf{r}^\mathcal{I}$. By definition of $p_{\mathsf{tp}_\mathcal{A}(\mathsf{a})}$ not being involved in any violation, we have $(p_{\mathsf{tp}_\mathcal{A}(\mathsf{a})}, e) \in \mathsf{s}^\mathcal{I}$. Then the definition of $\mathsf{s}^\mathcal{J}$ ensures $(\mathsf{a}, \rho(e)) \in \mathsf{s}^\mathcal{J}$, which again contradicts our assumption that $(\mathsf{a},d) \in \mathsf{r}^\mathcal{J} \setminus \mathsf{s}^\mathcal{J}$.

For Point 4, with the above observation that $\mathcal{J}$ equals the cost of $\mathcal{J}|_{\mathsf{crit}(\mathcal{I}) \cup W}$, it thus suffices to prove that the cost of $\mathcal{J}|_{\mathsf{crit}(\mathcal{I}) \cup W}$ is $\leq k$. To achieve this, we associate each violation in $\mathcal{J}$ to a corresponding violation in $\mathcal{I}$ in an injective manner. Unary violations in $\mathcal{J}|_{\mathsf{crit}(\mathcal{I})}$ are identical to those in $\mathcal{I}$, by Lemma 18. A unary violation on an element $w_t \in W$ can only be a concept inclusion violation, and thus is present on any element d from $\mathcal{I}$ such that $\mathsf{tp}_\mathcal{I}(d) = t$. Binary violations in $\mathcal{J}|_{\mathsf{crit}(\mathcal{I})} \times \mathcal{J}|_{\mathsf{crit}(\mathcal{I})}$ are identical to those in $\mathcal{I}$, by construction of $\mathcal{J}$. A binary violation on $(w_{t_1}, w_{t_2}) \in W \times W$ can only be a role inclusion $\mathsf{r} \sqsubseteq \mathsf{s}$ violation. By definition of $\mathsf{r}^\mathcal{J}$, there exists $(d_1,d_2) \in \mathsf{r}^\mathcal{I}$ with $d_1, d_2 \notin \mathsf{crit}(\mathcal{I})$, $\mathsf{tp}_\mathcal{I}(d_1) = t_1$ and $\mathsf{tp}_\mathcal{I}(d_2) = t_2$. It follows that $(d_1,d_2) \notin \mathsf{s}^\mathcal{I}$, which provides the desired witness for the role inclusion being violated already in $\mathcal{I}$. We now turn to the case of a binary violation on $(\mathsf{a}, w_t) \in \mathsf{crit}(\mathcal{I}) \times W$. First notice that it can again only be a violation of a role inclusion $\mathsf{r} \sqsubseteq \mathsf{s}$. By construction of $\mathsf{r}^\mathcal{J}$, there exists $d \in \Delta^\mathcal{I} \setminus \mathsf{crit}(\mathcal{I})$ such that $(\mathsf{a},d) \in \mathsf{r}^\mathcal{I}$ and $\mathsf{tp}_\mathcal{I}(d) = t$. It follows that $(\mathsf{a},d) \notin \mathsf{s}^\mathcal{I}$, which provides the desired witness for the role inclusion being violated already in $\mathcal{I}$. The case of a binary violation in $W \times \mathsf{crit}(\mathcal{I})$ is symmetric.

Point 5_p is immediate from $\rho(\mathcal{I}) \subseteq \mathcal{J}$ and the fact that ρ preserves the interpretation of all individuals occurring in q as $\mathsf{Ind}(q) \subseteq \mathsf{pc}(\mathcal{A}) \subseteq \mathsf{crit}(\mathcal{I})$. Point 5_c is a direct consequence of Lemma 18 and of $\mathsf{Ind}(q) \subseteq \mathsf{pc}(\mathcal{A})$. □

Proofs for Section 6.2 (Construction of the FO-Rewriting)

In what follows, we first clarify the full construction of the rewritten query, and then prove the correctness of the rewriting, stated in Lemma 9, to conclude the proof of Theorem 8.

For a given p-strategy (resp. c-strategy) $\sigma := (\mathcal{M}, \Gamma, \mathcal{V}_\nu)$, we denote by $\mathcal{S}_\sigma$ the set of *p-safe* ABox types for σ. As already outlined in the main body of the paper, the rewritten FO-query q' is obtained as the disjunction of queries q_σ, where each σ is a p-strategy (resp. c-strategy) for $(\mathcal{T}_{\omega_\mathcal{T}}, q, k)$. Now, for a given $\sigma := (\mathcal{M}, \Gamma, \mathcal{V}_\nu)$, the subquery q_σ uses one variable v_d for each $d \in \Gamma$; we use $\mathbf{v}_\Gamma$ to refer to the tuple formed by all these variables and V_Γ to refer to the set $\{v_d \mid d \in \Gamma\}$. Recall that if the query q_σ is to be satisfied, it must identify a subpart of the input ABox that could be interpreted as $\mathcal{M}|_\Gamma$, up to some violations described in $\mathcal{V}_\nu$. We describe the different subqueries which are used to define q_σ. First, we must verify that, in this pinpointed part of the ABox, every concept assertion that is unsatisfied in $\mathcal{M}$ is a violation from $\mathcal{V}$ with the expected cost. For a given concept name $\mathsf{A} \in \mathsf{N_C}(\mathcal{T})$ and $d \in \Gamma$, we let $q_{\sigma,\mathsf{A},d}(v_d)$ be the following formula:

$$\mathsf{A}(v_d) \wedge d \notin \mathsf{A}^\mathcal{M} \rightarrow \mathsf{A}(d) \in \mathcal{V} \wedge \mathsf{W}_\mathsf{A}^{\nu(\mathsf{A}(d))}(v_d),$$

where $d \notin \mathsf{A}^\mathcal{M}$ and $\mathsf{A}(d) \in \mathcal{V}$ are Boolean expressions that are replaced by truth constants $\top$ or $\bot$ in the above depending on σ. We shall use similar expressions in later subqueries without further mention. We proceed similarly for role assertions: for a given role name $\mathsf{r} \in \mathsf{N_R}(\mathcal{T})$ and $d, e \in \Gamma$, we let $q_{\sigma,\mathsf{r},d,e}(v_d, v_e)$ be the following formula:

$$\mathsf{r}(v_d, v_e) \wedge (d,e) \notin \mathsf{r}^\mathcal{M} \rightarrow \mathsf{r}(d,e) \in \mathcal{V} \wedge \mathsf{w}_\mathsf{r}^{\nu(\mathsf{r}(d,e))}(v_d, v_e).$$

We also make sure that the other individuals from the input ABox have an ABox type that is p-safe (resp. c-safe) w.r.t. σ. It is clear that 'variable v maps onto an individual with ABox type t' can be captured as follows:

$$q_t(v) := \bigwedge_{\mathsf{A} \in \mathsf{N_C}(\mathcal{T})} \mathsf{A}(v) \leftrightarrow \mathsf{A} \in t \quad \wedge \quad \bigwedge_{\mathsf{r} \in \mathsf{N_R^\pm}(\mathcal{T})} (\exists v'\ \mathsf{r}(v,v')) \leftrightarrow \exists \mathsf{r} \in t \quad \wedge \quad \bigwedge_{\mathsf{r} \in \mathsf{N_R^\pm}(\mathcal{T})} (\exists_{>k} v'\ \mathsf{r}(v,v')) \leftrightarrow \exists^{>k} \mathsf{r} \in t,$$

where $\exists_{>k} v\ \phi(v)$ is the usual counting quantifier. It remains to check whether the minimal completion of each role assertion involving an individual in $\mathcal{M}$ and one not involved in $\mathcal{M}$ is consistent w.r.t. $\mathcal{M}$. For $\mathsf{r} \in \mathsf{N_R^\pm}(\mathcal{T})$ and $d \in \Delta^\mathcal{M}$, we define $q_{\sigma,\mathsf{r},d}(\mathbf{v}_\Gamma)$ as the following formula:

$$\forall u \big(\mathsf{r}(v_d, u) \wedge \bigwedge_{e \in \Gamma} u \neq v_e \rightarrow \bigwedge_{\mathsf{r} \sqsubseteq_\mathcal{T} \mathsf{s}} \exists \mathsf{s} \in \mathsf{tp}_\mathcal{M}(d) \big)$$

We now combine these subqueries to obtain the query q_σ:

$$q_\sigma := \exists \mathbf{v}_\Gamma \Big(\bigwedge_{\substack{d,e \in \Gamma \\ d \neq e}} v_d \neq v_e \wedge \bigwedge_{\substack{\mathsf{A} \in \mathsf{N_C} \\ d \in \Gamma}} q_{\sigma,\mathsf{A},d}(v_d) \wedge \bigwedge_{\substack{\mathsf{r} \in \mathsf{N_R} \\ d,e \in \Gamma}} q_{\sigma,\mathsf{r},d,e}(v_d, v_e) \wedge \forall v \left(\bigwedge_{d \in \Gamma} v \neq v_d \rightarrow \bigvee_{t \in \mathcal{S}_\sigma} q_t(v) \right) \wedge \bigwedge_{\substack{\mathsf{r} \in \mathsf{N_R} \\ d \in \Gamma}} q_{\sigma,\mathsf{r},d}(\mathbf{v}_\Gamma) \Big)$$

Notice the first conjunction of inequalities simply enforces the SNA on ABox individuals. Now, as announced, we define:

$$q'_p := \bigvee_{\substack{\sigma \text{ is a p-strategy} \\ \text{for } (\mathcal{T},\omega_\mathcal{T},k,q)}} q_\sigma \qquad q'_c := \bigvee_{\substack{\sigma \text{ is a c-strategy} \\ \text{for } (\mathcal{T},\omega_\mathcal{T},k,q)}} q_\sigma.$$

Lemma 9. *Let $\mathcal{T}_{\omega_\mathcal{T}}$ be a weighted* DL-Lite$^{\mathcal{H}}_{\text{bool}}$ *TBox, k an integer, and q a BCQ. For every weighted ABox $\mathcal{A}_{\omega_\mathcal{A}}$:*

$$(\mathcal{T},\mathcal{A})_{\omega_\mathcal{T}\cup\omega_\mathcal{A}} \models^k_p q \text{ iff } \mathcal{I}_{\mathcal{A}_{\omega_\mathcal{A}}} \models q'_p.$$

Furthermore, if q is an IQ, then:

$$(\mathcal{T},\mathcal{A})_{\omega_\mathcal{T}\cup\omega_\mathcal{A}} \not\models^k_c q \text{ iff } \mathcal{I}_{\mathcal{A}_{\omega_\mathcal{A}}} \models q'_c.$$

Proof. Most of the proof is common to both q'_p and q'_c, typically when it comes down to building an interpretation with cost $\leq k$. We only differentiate the two cases when it is about verifying query (non-)entailment. We voluntarily write q' indifferently, and consider strategies without specifying whether they are p-strategies or c-strategies.

($\Leftarrow$). Assume $\mathcal{A}^{\omega_\mathcal{A}} \models q'$. Therefore, there exists a strategy $\sigma := (\mathcal{M},\Gamma,\mathcal{V}_\nu)$ such that $\mathcal{A}^{\omega_\mathcal{A}} \models q_\sigma$. We denote $W := \Delta^\mathcal{M} \setminus \Gamma$. By definition of $\mathcal{A}^{\omega_\mathcal{A}} \models q_\sigma$, there exists a mapping $h_0 : V_\Gamma \mapsto \mathsf{Ind}(\mathcal{A})$ that makes the rest of q_σ true on $\mathcal{A}^{\omega_\mathcal{A}}$. In particular, from the first conjunct in q_σ, h_0 is injective. We set:

$$\begin{array}{rcl} h : \Delta^\mathcal{M} & \to & h_0(V_\Gamma) \cup W \\ d & \mapsto & \begin{cases} h_0(v_d) & \text{if } d \in \Gamma \\ d & \text{otherwise} \end{cases} \end{array}$$

Notice that h is injective, since h_0 is, and is thus a bijection. For each t in $\mathcal{S}_\sigma$, we choose a 1-type t' whose existence is guaranteed by the definition of t being safe, and we denote $\rho : t \mapsto t'$ this choice function over $\mathcal{S}_\sigma$. By Point 5 of the definition of safe types, if $\exists \mathsf{r} \in \bigcup_{t\in\mathcal{S}_\sigma} \rho(t)$, then there exists $d \in \Delta^\mathcal{M}$ such that $d \in \exists \mathsf{s}^-$ for every $\mathsf{s} \in \mathsf{N}^\pm_\mathsf{R}$ such that $\mathsf{r} \sqsubseteq_\mathcal{T} \mathsf{s}$. We choose one such element d_r per $\exists \mathsf{r} \in \bigcup_{t\in\mathcal{S}_\sigma} \rho(t)$ and denote $w_\mathsf{r} := h(d_\mathsf{r})$.

We now set $\mathcal{I}$ as the interpretation with domain $\mathsf{Ind}(\mathcal{A}) \cup W$ and that interprets every concept name A and role name r as:

$$\begin{aligned} \mathsf{A}^\mathcal{I} &:= h(\mathsf{A}^\mathcal{M}) \cup \{\mathsf{a} \in \mathsf{Ind}(\mathcal{A}) \setminus h_0(V_\Gamma) \mid \mathsf{A} \in \rho(\mathsf{tp}_\mathcal{A}(\mathsf{a}))\} \\ \mathsf{r}^\mathcal{I} &:= h(\mathsf{r}^\mathcal{M}) \\ &\quad \cup \left\{ (\mathsf{a},\mathsf{b}) \in \mathsf{Ind}(\mathcal{A})^2 \setminus h_0(V_\Gamma)^2 \,\middle|\, \begin{array}{l} \mathsf{s}(\mathsf{a},\mathsf{b}) \in \mathcal{A} \text{ for} \\ \text{some } \mathsf{s} \sqsubseteq_\mathcal{T} \mathsf{r} \end{array} \right\} \\ &\quad \cup \left\{ (\mathsf{b},\mathsf{a}) \in \mathsf{Ind}(\mathcal{A})^2 \setminus h_0(V_\Gamma)^2 \,\middle|\, \begin{array}{l} \mathsf{s}(\mathsf{a},\mathsf{b}) \in \mathcal{A} \text{ for} \\ \text{some } \mathsf{s} \sqsubseteq_\mathcal{T} \mathsf{r}^- \end{array} \right\} \\ &\quad \cup \{(\mathsf{a}, w_\mathsf{s}) \mid \mathsf{a} \in \mathsf{Ind}(\mathcal{A}) \setminus h_0(V_\Gamma), \exists \mathsf{s} \in \rho(\mathsf{tp}_\mathcal{A}(\mathsf{a})), \mathsf{s} \sqsubseteq_\mathcal{T} \mathsf{r}\} \\ &\quad \cup \{(w_\mathsf{s}, \mathsf{a}) \mid \mathsf{a} \in \mathsf{Ind}(\mathcal{A}) \setminus h_0(V_\Gamma), \exists \mathsf{s} \in \rho(\mathsf{tp}_\mathcal{A}(\mathsf{a})), \mathsf{s} \sqsubseteq_\mathcal{T} \mathsf{r}^-\} \end{aligned}$$

Notice that the two last lines are well defined as q_σ being satisfied guarantees that every $\mathsf{a} \in \mathsf{Ind}(\mathcal{A}) \setminus h_0(V_\Gamma)$ has a safe ABox type in $\mathcal{A}$, thus ρ is indeed defined on $\mathsf{tp}_\mathcal{A}(\mathsf{a})$.

Before verifying that $\mathcal{I}$ has the desired properties, we first establish the following lemma:

Lemma 19. *For every $d \in \Delta^\mathcal{M}$, we have $\mathsf{tp}_\mathcal{I}(h(d)) = \mathsf{tp}_\mathcal{M}(d)$. For every $\mathsf{a} \in \mathsf{Ind}(\mathcal{A}) \setminus h_0(V_\Gamma)$, we have $\mathsf{tp}_\mathcal{I}(\mathsf{a}) = \rho(\mathsf{tp}_\mathcal{A}(\mathsf{a}))$.*

Proof. Let $d \in \Delta^\mathcal{M}$. The inclusion $\mathsf{tp}_\mathcal{M}(d) \subseteq \mathsf{tp}_\mathcal{I}(h(d))$ is trivial since $h(\mathcal{M}) \subseteq \mathcal{I}$. We turn to the inclusion $\mathsf{tp}_\mathcal{I}(h(d)) \subseteq \mathsf{tp}_\mathcal{M}(d)$. Consider a concept name $\mathsf{A} \in \mathsf{tp}_\mathcal{I}(h(d))$. If $h(d) \in h_0(V_\Gamma)$, then by definition of $\mathsf{A}^\mathcal{I}$, we have some $e \in \Delta^\mathcal{M}$ such that $h(e) = h(d)$ and $e \in \mathsf{A}^\mathcal{M}$. Recall that h is injective, thus $d = e$. It follows that $\mathsf{A} \in \mathsf{tp}_\mathcal{M}(d)$. Consider a role $\mathsf{r} \in \mathsf{N}^\pm_\mathsf{R}(\mathcal{T})$ such that $\exists \mathsf{r} \in \mathsf{tp}_\mathcal{I}(h(d))$, that is we have some $e \in \Delta^\mathcal{I}$ such that $(h(d), e) \in \mathsf{r}^\mathcal{I}$. We distinguish based on the definition of $\mathsf{r}^\mathcal{I}$.

- If $(h(d), e) \in h(\mathsf{r}^\mathcal{M})$, then we are done.
- If $(h(d), e) \in \mathsf{Ind}(\mathcal{A})^2 \setminus h_0(V_\Gamma)^2$ and $\mathsf{s}(h(d), e) \in \mathcal{A}$ for some $\mathsf{s} \sqsubseteq_\mathcal{T} \mathsf{r}$, then $h(d) \in h_0(V_\Gamma)$ and $e \in \mathsf{Ind}(\mathcal{A}) \setminus h_0(V_\Gamma)$. Thus from $q_{\sigma,\mathsf{s},d}$ being true for $u := e$, we obtain $\exists \mathsf{r} \in \mathsf{tp}_\mathcal{M}(d)$ as desired.
- If $(e, h(d)) \in \mathsf{Ind}(\mathcal{A})^2 \setminus h_0(V_\Gamma)^2$ and $\mathsf{s}(h(d), e) \in \mathcal{A}$ for some $\mathsf{s} \sqsubseteq_\mathcal{T} \mathsf{r}^-$, then $h(d) \in h_0(V_\Gamma)$ and $e \in \mathsf{Ind}(\mathcal{A}) \setminus h_0(V_\Gamma)$. Thus from $q_{\sigma,\mathsf{s}^-,d}$ being true for $u := e$, we again obtain $\exists \mathsf{r} \in \mathsf{tp}_\mathcal{M}(d)$ as desired.
- Note that $h(d) \in \mathsf{Ind}(\mathcal{A}) \setminus h_0(V_\Gamma)$ is impossible.
- Otherwise $h(d) = w_\mathsf{s}$ for some $\mathsf{s} \sqsubseteq_\mathcal{T} \mathsf{r}^-$ such that there is $\mathsf{a} \in \mathsf{Ind}(\mathcal{A}) \setminus h_0(V_\Gamma)$ with $\exists \mathsf{s} \in \rho(\mathsf{tp}_\mathcal{A}(\mathsf{a}))$. From the definition of w_s we obtain $\exists \mathsf{t}^- \in \mathsf{tp}_\mathcal{M}(d)$ for every $\mathsf{s} \sqsubseteq_\mathcal{T} \mathsf{t}$. In particular for $\mathsf{t} := \mathsf{r}^-$, we get $\exists \mathsf{r} \in \mathsf{tp}_\mathcal{M}(d)$.

In all the above cases, we obtained $\exists \mathsf{r} \in \mathsf{tp}_{\mathcal{M}}(d)$, thus $\mathsf{tp}_{\mathcal{I}}(h(d)) \subseteq \mathsf{tp}_{\mathcal{M}}(d)$ which concludes this part of the proof.

Let $\mathsf{a} \in \mathsf{Ind}(\mathcal{A}) \setminus h_0(V_\Gamma)$. The inclusion $\rho(\mathsf{tp}_{\mathcal{A}}(\mathsf{a})) \subseteq \mathsf{tp}_{\mathcal{I}}(\mathsf{a})$ is trivial by construction of $\mathcal{I}$ (notice that, for the case of $\exists \mathsf{r} \in \rho(\mathsf{tp}_{\mathcal{A}}(\mathsf{a}))$, the interpretation provides the r-edge $(\mathsf{a}, w_{\mathsf{r}})$). We turn to the inclusion $\mathsf{tp}_{\mathcal{I}}(\mathsf{a}) \subseteq \rho(\mathsf{tp}_{\mathcal{A}}(\mathsf{a}))$. Consider a concept name $\mathsf{A} \in \mathsf{tp}_{\mathcal{I}}(\mathsf{a})$. Then by definition of $\mathsf{A}^{\mathcal{I}}$, we have $\mathsf{A} \in \rho(\mathsf{tp}_{\mathcal{A}}(\mathsf{a}))$ as desired. Consider a role $\mathsf{r} \in \mathsf{N}_{\mathsf{R}}^{\pm}(\mathcal{T})$ such that $\exists \mathsf{r} \in \mathsf{tp}_{\mathcal{I}}(\mathsf{a})$, that is, we have some $e \in \Delta^{\mathcal{I}}$ such that $(\mathsf{a}, e) \in \mathsf{r}^{\mathcal{I}}$. We distinguish based on the definition of $\mathsf{r}^{\mathcal{I}}$.

- From $\mathsf{a} \notin h_0(V_\Gamma)$, we cannot have $(\mathsf{a}, e) \in h(\mathcal{M})$.
- If $(\mathsf{a}, e) \in \mathsf{Ind}(\mathcal{A})^2 \setminus h_0(V_\Gamma)^2$ and $\mathsf{s}(\mathsf{a}, e) \in \mathcal{A}$ for some $\mathsf{s} \sqsubseteq_{\mathcal{T}} \mathsf{r}$, then $\exists \mathsf{s} \in \mathsf{tp}_{\mathcal{A}}(\mathsf{a})$. Since $\mathsf{tp}_{\mathcal{A}}(\mathsf{a})$ is a safe type (from q_σ being satisfied), Points 2 and 4 in the definition of safe-types, guarantees that $\exists \mathsf{r} \in \rho(\mathsf{tp}_{\mathcal{A}}(\mathsf{a}))$ as desired.
- If $(e, \mathsf{a}) \in \mathsf{Ind}(\mathcal{A})^2 \setminus h_0(V_\Gamma)^2$ and $\mathsf{s}(e, \mathsf{a}) \in \mathcal{A}$ for some $\mathsf{s} \sqsubseteq_{\mathcal{T}} \mathsf{r}^-$, then $\exists \mathsf{s}^- \in \mathsf{tp}_{\mathcal{A}}(\mathsf{a})$. Since $\mathsf{tp}_{\mathcal{A}}(\mathsf{a})$ is a safe type (from q_σ being satisfied), Points 2 and 4 in the definition of safe-types, guarantees that $\exists \mathsf{r} \in \rho(\mathsf{tp}_{\mathcal{A}}(\mathsf{a}))$ as desired.
- If $e = w_{\mathsf{s}}$ for some $\mathsf{s} \sqsubseteq_{\mathcal{T}} \mathsf{r}$ such that $\exists \mathsf{s} \in \rho(\mathsf{tp}_{\mathcal{A}}(\mathsf{a}))$, then Point 4 in the definition of safe types yields $\exists \mathsf{r} \in \rho(\mathsf{tp}_{\mathcal{A}}(\mathsf{a}))$.
- Note that a cannot be a witness w_{s} as those are not individuals.

This concludes the proof of $\mathsf{tp}_{\mathcal{I}}(\mathsf{a}) \subseteq \rho(\mathsf{tp}_{\mathcal{A}}(\mathsf{a}))$. □

We now verify that $\mathcal{I}$ is an interpretation with cost $\leq k$. First consider a concept assertion $\mathsf{C}(\mathsf{a}) \in \mathcal{A}$ violated in $\mathcal{I}$. If $\mathsf{a} \in \mathsf{Ind}(\mathcal{A}) \setminus h_0(V_\Gamma)$, then $\mathsf{tp}_{\mathcal{A}}(\mathsf{a})$ is a safe type, from q_σ being satisfied, and thus $\mathsf{a} \in \mathsf{A}^{\mathcal{I}}$ by definition of $\mathsf{A}^{\mathcal{I}}$ and Point 1 in the definition of safe types; this contradicts $\mathsf{C}(\mathsf{a})$ being violated in $\mathcal{I}$. Otherwise $\mathsf{a} \in h_0(V_\Gamma)$, so there exists $d \in \Gamma$ such that $h_0(v_d) = \mathsf{a}$. From $\mathsf{A}(\mathsf{a})$ being violated, we have $d \notin \mathsf{A}^{\mathcal{M}}$, thus from $q_{\sigma,\mathsf{A},d}$ being satisfied we obtain $\mathsf{A}(\mathsf{d}) \in \mathcal{V}$ and $\mathsf{W}_{\mathsf{A}}^{\nu(\mathsf{A}(d))}(\mathsf{a})$, thus the weight of $\mathsf{A}(\mathsf{a})$ is accounted for in $\mathcal{V}_\nu$.

We proceed similarly for a role assertion $\mathsf{r}(\mathsf{a}, \mathsf{b}) \in \mathcal{A}$. It is satisfied if $(\mathsf{a}, \mathsf{b}) \in \mathsf{Ind}(\mathcal{A})^2 \setminus h_0(V_\Gamma)^2$ by definition of $\mathsf{r}^{\mathcal{I}}$. If $(\mathsf{a}, \mathsf{b}) \in h_0(V_\Gamma)^2$, then the cost is accounted for in $\mathcal{V}_\nu$, this time using the corresponding subquery $q_{\sigma,\mathsf{r},h^{-1}(\mathsf{a}),h^{-1}(\mathsf{b})}$. Therefore, the cost of ABox violations in $\mathcal{I}$ is bounded by $\sum_{\alpha \in \mathcal{V}} \nu(\alpha)$.

The case of concept inclusions is rather straightforward with Lemma 19 at hand. If a violation of $\mathsf{C} \sqsubseteq \mathsf{D}$ occurs at $d \in \Delta^{\mathcal{I}}$, then we distinguish two cases. If $d \in h(\Delta^{\mathcal{M}})$, then Lemma 19 guarantees that $h^{-1}(d)$ also violates $\mathsf{C} \sqsubseteq \mathsf{D}$ in $\mathcal{M}$, thus the cost of this violation is accounted for in the TBox violations of $\mathcal{M}$. If $d \notin h(\Delta^{\mathcal{M}})$, then Lemma 19 yields $\mathsf{tp}_{\mathcal{I}}(d) = \rho(\mathsf{tp}_{\mathcal{A}}(d))$. The latter, using Point 3 from the definition of safe types, does not violate any CIs, which contradicts d violating $\mathsf{C} \sqsubseteq \mathsf{D}$.

For a role inclusion $\mathsf{r} \sqsubseteq \mathsf{s}$, we trivially have $\mathsf{r} \sqsubseteq_{\mathcal{T}} \mathsf{s}$. Therefore, in the definition of $\mathsf{r}^{\mathcal{I}}$, it is clear that only the case of $h(\mathsf{r}^{\mathcal{M}})$ may be violated (that is, not-matched with a pair from $h(\mathsf{s}^{\mathcal{M}})$), as the remaining cases are immediately treated by $\sqsubseteq_{\mathcal{T}}$ being transitive. However, every violation of a role inclusion in $h(\mathcal{M})$ is a violation of the same role inclusion in $\mathcal{M}$, thus its cost is already accounted for in the TBox violations of $\mathcal{M}$.

Overall, we obtained that $(\omega_{\mathcal{T}} \cup \omega_{\mathcal{A}})(\mathcal{I}) \leq \omega_{\mathcal{T}}(\mathcal{M}) + \sum_{\alpha \in \mathcal{V}} \nu(\alpha)$. Recalling the definition of a strategy, we obtain $(\omega_{\mathcal{T}} \cup \omega_{\mathcal{A}})(\mathcal{I}) \leq k$. It remains to verify that $\mathcal{I}$ behaves as expected w.r.t. the query q.

- If working with q a CQ under the possible semantics, then the definition of p-strategies guarantees that $\mathcal{M} \models q$. Therefore, from $h(\mathcal{M}) \subseteq \mathcal{I}$ and $\mathsf{Ind}(q) \subseteq \Gamma \subseteq \Delta^{\mathcal{M}}$, we obtain immediately that $\mathcal{I} \models q$ as desired.
- If working with q an IQ under the certain semantics, then the definition of c-strategies guarantees that $\mathsf{Ind}(q) \subseteq \Gamma$ and $\mathcal{M} \not\models q$. Recall that we restricted w.l.o.g. to the case of q being a concept IQ, we thus distinguish between two possible cases. If $q = \mathsf{A}(\mathsf{a})$, then Lemma 19 joint with the above directly concludes. If $q = \exists y\, \mathsf{A}(y)$, then we use Lemma 19 to obtain $\mathsf{A}^{\mathcal{I}} \cap h(\Delta^{\mathcal{M}}) = \emptyset$, and use Point 6 in the definition of a c-safe type to also ensure $\mathsf{A}^{\mathcal{I}} \cap (\Delta^{\mathcal{M}} \setminus h(\Delta^{\mathcal{M}})) = \emptyset$.

($\Rightarrow$). Assume that there exists an interpretation $\mathcal{I}$ of $(\mathcal{T}, \mathcal{A})_{\omega_{\mathcal{T}} \cup \omega_{\mathcal{A}}}$ whose cost is $\leq k$ and that either satisfies the CQ q (for the possible semantics), or does not satisfy the IQ q (for the certain semantics). We apply Lemma 8 and denote $\mathcal{J}$ the obtained interpretation whose domain is $\mathsf{Ind}(\mathcal{A}) \cup W$ for some set of 1-types W. Let $\mathcal{M} := \mathcal{J}|_{\mathsf{core}(\mathcal{A}) \cup W}$, $\Gamma := \mathsf{core}(\mathcal{A})$, $\mathcal{V} := \{\alpha \in \mathcal{A} \mid \mathcal{J} \not\models \alpha\}$ and $\nu : \alpha \mapsto \omega_{\mathcal{A}}(\alpha)$ the cost function over $\mathcal{V}$. Denote $\sigma := (\mathcal{M}, \Gamma, \mathcal{V}_\nu)$ the corresponding strategy. It is direct to verify that σ is indeed a p-strategy, resp. a c-strategy, depending on the considered case. We claim that $\mathcal{A}^\omega \models q_\sigma$, and thus $\mathcal{A}^\omega \models q'$. It is not hard to verify it, by considering the following mapping h that defines how to assign the existentially quantified variables $\mathbf{v}_\Gamma$ of q_σ:

$$\begin{array}{rcl} h : V_\Gamma & \to & \Gamma \\ v_d & \mapsto & d. \end{array}$$

We highlight two key ingredients to verify the above claim:

- To check that the part of q_σ interested in safe types is satisfied, one needs to exhibit fitting 1-types t' for an ABox type t found on some individual a that is not reached by the variables $\mathbf{v}_\Gamma$. This is achieved by setting $t' := \mathsf{tp}_{\mathcal{I}}(p_{\mathsf{tp}_{\mathcal{A}}(\mathsf{a})})$. From Point 3 of Lemma 8, guaranteeing that non-core elements are not involved in any violations, it is clear that t' satisfies Points 1-4 from the definition of safe types. To check that t' satisfies Point 5 follows from Lemma 18 and the construction of $\mathcal{J}$. Points 6 and 7 (in the case of c-safe types) both follow from Point 5_c of Lemma 8.
- Using Lemma 18, one can verify that such t' can always be found already in $\mathcal{J}$. To check that each subquery $q_{\sigma,\mathsf{r},d}$ is indeed satisfied, we rely again on Point 3 from Lemma 8 that guarantees non-core elements are not involved in any violations. □