Data-driven inverse optimization with imperfect information

Mohajerin Esfahani, P.; Shafieezadeh-Abadeh, Soroosh; Hanasusanto, Grani A.; Kuhn, Daniel

DOI
10.1007/s10107-017-1216-6

Publication date
2018

Document Version
Final published version

Published in
Mathematical Programming

Math. Program., Ser. B (2018) 167:191–234
https://doi.org/10.1007/s10107-017-1216-6

FULL LENGTH PAPER

Data-driven inverse optimization with imperfect information

Peyman Mohajerin Esfahani[1] · **Soroosh Shafieezadeh-Abadeh**[2] · **Grani A. Hanasusanto**[3] · **Daniel Kuhn**[2]

Received: 11 December 2015 / Accepted: 28 November 2017 / Published online: 7 December 2017

Abstract In data-driven inverse optimization an observer aims to learn the preferences of an agent who solves a parametric optimization problem depending on an exogenous signal. Thus, the observer seeks the agent's objective function that best explains a historical sequence of signals and corresponding optimal actions. We focus here on situations where the observer has imperfect information, that is, where the agent's true objective function is not contained in the search space of candidate objectives, where the agent suffers from bounded rationality or implementation errors, or where the observed signal-response pairs are corrupted by measurement noise. We formalize this inverse optimization problem as a distributionally robust program minimizing the worst-case risk that the *predicted* decision (i.e., the decision implied by a particular candidate objective) differs from the agent's *actual* response to a random signal. We show that our framework offers rigorous out-of-sample guarantees for different loss functions used to measure prediction errors and that the emerging inverse optimization problems can be exactly reformulated as (or safely approximated by) tractable

✉ Daniel Kuhn
daniel.kuhn@epfl.ch

Peyman Mohajerin Esfahani
P.MohajerinEsfahani@tudelft.nl

Soroosh Shafieezadeh-Abadeh
soroosh.shafiee@epfl.ch

Grani A. Hanasusanto
grani.hanasusanto@utexas.edu

1 Delft Center for Systems and Control, TU Delft, Delft, The Netherlands

2 Risk Analytics and Optimization Chair, EPFL, Lausanne, Switzerland

3 Graduate Program in Operations Research and Industrial Engineering, UT Austin, Austin, TX, USA

convex programs when a new suboptimality loss function is used. We show through extensive numerical tests that the proposed distributionally robust approach to inverse optimization attains often better out-of-sample performance than the state-of-the-art approaches.

Mathematics Subject Classification C15 Stochastic programming · 90C25 Convex programming · 90C47 Minimax problems

1 Introduction

In inverse optimization an observer aims to learn the preferences of an agent who solves a parametric optimization problem depending on an exogenous signal. The observer knows the constraints imposed on the agent's actions but is unaware of her objective function. By monitoring a sequence of signals and corresponding actions, the observer seeks to identify an objective function that makes the observed actions optimal in the agent's optimization problem. This learning problem can be cast as an *inverse optimization problem* over candidate objective functions. The hope is that the solution of this inverse problem enables the observer to predict the agent's future actions in response to new signals.

Inverse optimization has a wide spectrum of applications spanning several disciplines ranging from econometrics and operations research to engineering and biology. For example, a marketing executive aims to understand the purchasing behavior of consumers with unknown utility functions by monitoring sales figures [1,7,11], a transportation planner wishes to learn the route choice preferences of the passengers in a multimodal transport system by measuring traffic flows [3,15,17,20], or a healthcare manager seeks to design clinically acceptable treatments in view of historical treatment plans [19]. It is even believed that the behavior of many biological systems is governed by a principle of optimality with respect to an unknown decision criterion, which can be inferred by tracking the system [13,39]. Inverse optimization has also been applied in geoscience [32,42], portfolio selection [10,26], production planning [40], inventory management [18], network design and control [3,17,20,25] and the analysis of electricity prices [34].

The main thrust of the early literature on inverse optimization is to identify an objective function that explains a *single* observation. In the seminal paper [4] the agent solves a static (non-parametric) linear program and reveals her optimal decision to the observer, who then identifies the objective function closest to a prescribed nominal objective, under which the observed decision is optimal. This model was later extended to conic programs [26], integer programs [2,24,35,41] and linearly constrained separable convex programs [43]. Another variant of this problem is considered in [2], where the observer identifies an admissible objective function for which the optimal value of the agent's problem is closest to the observed optimal value corresponding to the unknown true objective.

This paper focuses on *data-driven* inverse optimization problems where the agent solves a parametric optimization problem *several times*. Accordingly, the observer has access to a finite sequence of signals and corresponding optimal responses. Using

this training data, the observer aims to infer an objective function that accurately predicts the agent's optimal responses to unseen future signals. As in classical regression, this learning task could be addressed by minimizing an empirical loss that penalizes the mismatch between the predicted and true optimal responses to a given signal. *Data-driven* inverse optimization problems of this type have only just started to attract attention, and to the best of our knowledge there are currently only three papers that study such problems. In [27] the observer seeks an objective function under which all observed decisions solve the Karush–Kuhn–Tucker (KKT) optimality conditions of the agent's convex optimization problem. To this end, the observer minimizes some norm of the KKT residuals at all observations. A similar goal is pursued in [11], where the optimality conditions are expressed via variational inequalities that can be reformulated as tractable conic constraints using ideas from robust optimization. This approach has the additional benefit that it extends to more general inverse equilibrium problems, which indicates that inverse optimization problems constitute special instances of mathematical programs with equilibrium constraints. A comprehensive survey of variational inequalities and mathematical programs with equilibrium constraints is provided in [23]. The third paper suggests to minimize the empirical average of the squared Euclidean distances between the predicted and true observed decisions, in which case the data-driven inverse optimization problem reduces to a bilevel program [6].

In summary, all existing approaches to data-driven inverse optimization solve an empirical loss minimization problem over some search space of candidate objectives. Different approaches mainly differ with respect to the loss functions that capture the mismatch between predictions and observations. The *KKT loss* used in [27] quantifies the extent to which the observed response to some signal violates the KKT conditions for a fixed candidate objective. Similarly, the *first-order loss* used in [11] measures the extent to which an observed response violates the first-order optimality conditions. Moreover, the *predictability loss* used in [6] captures the squared distance between an observed response and the response predicted by a given candidate objective.

In this paper we introduce the new *suboptimality loss*, which quantifies the degree of suboptimality of an observed response under a given candidate objective. While the predictability and suboptimality losses both enjoy a direct physical meaning, the KKT and first-order losses are generally not as easily interpretable.

Computational experiments in [27] and [11] suggest that empirical loss minimization problems under perfect information are likely to correctly identify the agent's true objective function if there is sufficient training data and the search space of candidate objectives is not too large. In any realistic setting, however, the observer is confronted with imperfect information such as *model uncertainty* (the agent's true objective is not one of the candidate objectives), *noisy measurements* (the observed signals and responses are corrupted by measurement errors) or *bounded rationality* (the agent settles for suboptimal responses due to cognitive or computational limitations). Due to overfitting effects, imperfect information can severely impair the predictive power of a candidate objective obtained via empirical loss minimization. This is simply a manifestation of the notorious 'garbage in-garbage out' phenomenon. As imperfect information certainly reflects the norm rather than the exception in inverse optimization, we propose here a systematic approach to combat overfitting via distributionally robust optimization. Specifically, inspired by [30] and [37], we use the Wasserstein

distance to construct a ball in the space of all signal-response-distributions centered at the empirical distribution on the training samples, and we formulate a distributionally robust inverse optimization problem that minimizes the worst-case risk of loss for any combination of a risk measure with a loss function, where the worst case is taken over all distributions in the Wasserstein ball. If the radius of the Wasserstein ball is chosen judiciously, we can guarantee that it contains the unknown data-generating distribution with high confidence, which in turn allows us to derive rigorous out-of-sample guarantees for the risk of loss of unseen future observations. The proposed distributionally robust inverse optimization problem can naturally be interpreted as a regularization of the corresponding empirical loss minimization problem. While regularization is known to improve the out-of-sample performance of numerous estimators in statistics, it has not yet been investigated systematically in the context of data-driven inverse optimization.

We highlight the following main contributions of this paper relative to the existing literature:

- We propose the suboptimality loss as an alternative to the KKT, first-order and predictability losses. The suboptimality loss admits a direct physical interpretation (like the predictability loss) and leads to convex empirical loss minimization problems (like the KKT and first-order losses) whenever the candidate objective functions admit a linear parameterization. In contrast, empirical predictability loss minimzation problems constitute NP-hard bilevel programs even for linear candidate objectives. We also propose the bounded rationality loss, which generalizes the suboptimality loss to situations where the agent is known to select δ-suboptimal decision due to bounded rationality.
- We leverage the data-driven distributionally robust optimization scheme with Wasserstein balls developed in [30] to regularize empirical inverse optimization problems under imperfect information. As such, the proposed approach offers out-of-sample guarantees for any combination of risk measures and loss functions. In contrast, [11] develops out-of-sample guarantees only for the value-at-risk of the first-order loss, while [27] and [6] discuss no (finite) out-of-sample guarantees at all.
- We study the tractability properties of the distributionally robust inverse optimization problem that minimizes the conditional value-at-risk of the suboptimality loss. We prove that this problem is equivalent to a convex program when the search space consists of all linear functions. We also show that this problem admits a safe convex approximation when the search space consists of all convex quadratic functions or all conic combinations of finitely many convex basis functions.
- We argue that the first-order and suboptimality losses can be used as tractable approximations for the intractable predictability loss, which has desirable statistical consistency properties [6] and is the preferred loss function if the observer aims for prediction accuracy. We show that if the candidate objective functions are strongly convex, then the estimators obtained from minimizing the first-order and suboptimality losses admit out-of-sample predictability guarantees. Moreover, the predictability guarantee corresponding to the suboptimality loss is stronger than

the one obtained from the first-order loss. Recall that the predictability loss itself cannot be minimized in polynomial time.

- We show through extensive numerical tests that the proposed distributionally robust approach to inverse optimization attains often better (lower) out-of-sample suboptimality and predictability than the state-of-the art approaches in [11] and [6]. All of our experiments are reproducible, and the underlying source codes are available at https://github.com/sorooshafiee/InverseOptimization.

The rest of the paper develops as follows. In Sects. 2 and 3 we formalizes the inverse optimization problem under perfect and imperfect information, respectively. Section 4 then introduces the distributionally robust approach to inverse optimization, while Sects. 5 and 6 derive tractable reformulations and safe approximations for distributionally robust inverse optimization problems over search spaces of linear and quadratic candidate objectives, respectively. Numerical results are reported in Sect. 7.

Notation The inner product of two vectors $s, t \in \mathbb{R}^m$ is denoted by $\langle s, t \rangle := s^\intercal t$, and the dual of a norm $\|\cdot\|$ on $\mathbb{R}^m$ is defined through $\|t\|_* := \sup_{\|s\| \leq 1} \langle t, s \rangle$. The dual of a proper (closed, solid, pointed) convex cone $\mathcal{C} \subseteq \mathbb{R}^m$ is defined as $\mathcal{C}^* := \{t \in \mathbb{R}^m : \langle t, s \rangle \geq 0 \ \forall s \in \mathcal{C}\}$, and the relation $s \succeq_{\mathcal{C}} t$ is interpreted as $s - t \in \mathcal{C}$. Similarly, for two symmetric matrices $Q, R \in \mathbb{R}^{m \times m}$ the relation $Q \succeq R$ ($Q \preceq R$) means that $Q - R$ is positive (negative) semidefinite. The identity matrix is denoted by $\mathbb{I}$. We denote by δ_ξ the Dirac distribution concentrating unit mass at $\xi \in \Xi$. The N-fold product of a distribution $\mathbb{P}$ on Ξ is denoted by $\mathbb{P}^N$, which represents a distribution on the Cartesian product Ξ^N. The decision variables of an optimization model are always specified directly under the optimization operator or in the first few lines of the list of constraints.

2 Inverse optimization under perfect information

Consider an agent who first receives a random signal $s \in \mathbb{S} \subseteq \mathbb{R}^m$ and then solves the following parametric optimization problem:

$$\underset{x \in \mathbb{X}(s)}{\text{minimize}} \ F(s, x). \tag{1}$$

Note that both the objective function $F : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ as well as the (multivalued) feasible set mapping $\mathbb{X} : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ depend on the signal. We assume that the set of minimizers $\mathbb{X}^\star(s) := \arg\min_{x \in \mathbb{X}(s)} F(s, x)$ is non-empty for every $s \in \mathbb{S}$. Consider also an independent observer who monitors the signal $s \in \mathbb{S}$ as well as the agent's optimal response $x \in \mathbb{X}^\star(s)$. We assume that the observer is ignorant of the agent's preferences encoded by the objective function F. Thus, a priori, the observer cannot predict the agent's response x to a particular signal s. Throughout the paper we assume that the observed signal-response pairs $\xi := (s, x)$ are governed by some probability distribution $\mathbb{P}$ supported on $\Xi := \big\{(s, x) : s \in \mathbb{S}, \ x \in \mathbb{X}(s)\big\}$, which can be viewed as the graph of the feasible set mapping $\mathbb{X}$. Note that the marginal distribution of s under $\mathbb{P}$ captures the frequency of the exogenous signals, while the conditional distribution of x given s places all probability mass on the argmin set $\mathbb{X}^\star(s)$. Note that, unless

$\mathbb{X}^\star(s)$ is a singleton, the exact conditional distribution of x given s depends on the specific optimization algorithm used by the agent.

In the following we assume that the observer has access to N independent samples $\widehat{\xi}_i := (\widehat{s}_i, \widehat{x}_i)$ from $\mathbb{P}$, which can be used to learn the agent's objective function. As the space of all possible objective functions is vast, the observer seeks to approximate F by some candidate objective function from within a parametric *hypothesis space* $\mathcal{F} = \{F_\theta : \theta \in \Theta\}$, where Θ represents a finite-dimensional parameter set. Ideally, the observer would aim to identify the hypothesis F_θ closest to F, *e.g.*, by solving the least squares problem

$$\underset{\theta\in\Theta}{\text{minimize}} \ \frac{1}{N}\sum_{i=1}^{N} \ell_\theta(\widehat{s}_i, \widehat{x}_i), \tag{2}$$

where $\ell_\theta(\widehat{s}_i, \widehat{x}_i)$ denotes the identifiability loss as per the following definition.

Definition 2.1 (*Identifiability loss*) The identifiability loss of model θ is given by

$$\ell_\theta(s, x) := |F(s, x) - F_\theta(s, x)|^2. \tag{3}$$

Unfortunately, the identifiability loss of a training sample cannot be evaluated unless the agent's objective function F is known. Indeed, the observer is blind to the agent's objective values $F(\widehat{s}_i, \widehat{x}_i)$ and only sees the signals $\widehat{s}_i$ and responses $\widehat{x}_i$. Thus, the identifiability loss cannot be used to learn F. It can merely be used to assess the quality of a hypothesis F_θ obtained with another method in a synthetic experiment where the true objective F is known. As two objective functions have the same minimizers whenever they are related through a strictly monotonically increasing transformation, however, it is indeed fundamentally impossible to learn F from the available training data. At best we can learn the set of its minimizers $\mathbb{X}^\star(s)$ for every s.

If a hypothesis F_θ is used in lieu of F, it can be used to predict the agent's optimal response to a signal s by solving a variant of problem (1), where F is replaced with F_θ. In the following we define $\mathbb{X}^\star_\theta(s) := \arg\min_{y\in\mathbb{X}(s)} F_\theta(s, y)$ and refer to any $x \in \mathbb{X}^\star_\theta(s)$ as a response to s predicted by θ. Note that $x \in \mathbb{X}^\star_\theta(s)$ if and only if the response x to s can be explained by model θ. In order to assess the quality of a candidate model θ, the observer should now check whether $\mathbb{X}^\star_\theta(s) \approx \mathbb{X}^\star(s)$ with high probability over $s \in \mathbb{S}$. This can be achieved by solving an empirical loss minimization problem of the form (2) with a loss function that satisfies $\ell_\theta(s, x) = 0$ if $x \in \mathbb{X}^\star_\theta(s)$ and $\ell_\theta(s, x) > 0$ otherwise. Thus, the loss should vanish if and only if the decision x can be explained as an optimal response to s under model θ.

In order to learn $\mathbb{X}^\star(s)$, an intuitive approach is to minimize the predictability loss defined below.

Definition 2.2 (*Predictability loss*) The predictability loss of model θ is given by

$$\ell_\theta(s, x) := \min_{y\in\mathbb{X}^\star_\theta(s)} \|x - y\|_2^2. \tag{4a}$$

It quantifies the squared Euclidean distance of x from the set of responses to s predicted by θ.

The predictability loss is known to offer strong statistical consistency guarantees but renders (2) an NP-hard bilevel optimization problem even if the agent's subproblem is convex [6]. Thus, the predictability loss can only be used for low-dimensional problems involving moderate sample sizes. An alternative choice is to minimize the suboptimality loss proposed in this paper, which we will show to be computationally attractive.

Definition 2.3 (*Suboptimality loss*) The suboptimality loss of model θ is given by

$$\ell_\theta(s, x) := F_\theta(s, x) - \min_{y \in \mathbb{X}(s)} F_\theta(s, y). \tag{4b}$$

It quantifies the suboptimality of x, that is, the cost of x in excess to the minimum cost under F_θ given s.

Another computationally attractive loss function is the degree of violation of the agent's first-order optimality condition [11].

Definition 2.4 (*First-order loss*) If F_θ is differentiable with respect to x, the first-order loss is given by

$$\ell_\theta(s, x) := \max_{y \in \mathbb{X}(s)} \big\langle \nabla_x F_\theta(s, x), x - y \big\rangle. \tag{4c}$$

It quantifies the extent to which x violates the first-order optimality condition of the optimization problem (1) for a given s, where F is replaced with F_θ. Note that the first-order loss vanishes whenever x represents a local minimizer of $F_\theta(s, \cdot)$ over $\mathbb{X}(s)$.

Note that the predictability loss best captures the observer's objective to predict the agent's decisions. However, the suboptimality loss and the first-order loss have better computational properties. Indeed, we will argue below that the learning model (2) with the losfunctions (4b) or (4c) is computationally tractable under suitable convexity assumptions about the agent's decision problem (1), the support set Ξ and the hypothesis space $\mathcal{F}$. Thus, we encounter a similar situation as in binary classification, where it is preferable to minimize the convex hinge loss instead of the discontinuous 0-1 loss, which is the actual quantity of interest. We also emphasize that the first-order and suboptimality losses coincide for linear hypotheses; see Sect. 5.

The following proposition establishes basic properties of the loss functions (4).

Proposition 2.5 (Dominance relations between loss functions) *Assume that $F_\theta(s, x)$ is convex and differentiable in x, and define $\gamma \geq 0$ as the largest number satisfying the inequality*

$$F_\theta(s, y) - F_\theta(s, x) \geq \big\langle \nabla_x F_\theta(s, x), y - x \big\rangle + \frac{\gamma}{2} \|y - x\|^2 \quad \forall x, y \in \mathbb{X}(s), \ \forall s \in \mathbb{S}. \tag{5}$$

If ℓ_θ^{p}, ℓ_θ^{s} and ℓ_θ^{f} denote the predictability, suboptimality and first-order losses, respectively, then we have

$$\ell_\theta^{\mathrm{f}}(s,x) \;\geq\; \ell_\theta^{\mathrm{s}}(s,x) \;\geq\; \frac{\gamma}{2}\,\ell_\theta^{\mathrm{p}}(s,x) \qquad \forall s \in \mathbb{S},\; x \in \mathbb{X}(s). \tag{6}$$

Moreover, all three loss functions are non-negative and evaluate to zero if and only if $x \in \mathbb{X}_\theta(s)$.

Note that (5) always holds for $\gamma = 0$ due to the first-order condition of convexity [14, Section 3.1.3].

Proof of Proposition 2.5 Setting $\gamma = 0$ and minimizing both sides of (5) over $y \in \mathbb{X}(s)$ yields $\ell_\theta^{\mathrm{f}}(s,x) \geq \ell_\theta^{\mathrm{s}}(s,x)$. Next, the first-order optimality condition of the convex program (1) with objective function F_θ requires that

$$\big\langle \nabla_x F_\theta(s,x),\, y - x \big\rangle \geq 0 \qquad \forall y \in \mathbb{X}(s) \tag{7}$$

at any optimal point $x \in \mathbb{X}_\theta^\star(s)$. Combining the inequalities (5) and (7) then yields

$$F_\theta(s,y) - F_\theta(s,x) \geq \frac{\gamma}{2}\|y - x\|^2 \qquad \forall x \in \mathbb{X}_\theta^\star(s),\; y \in \mathbb{X}(s).$$

Minimizing both sides of the above inequality over $x \in \mathbb{X}_\theta^\star(s)$ yields $\ell_\theta^{\mathrm{s}}(s,x) \geq \frac{\gamma}{2}\ell_\theta^{\mathrm{p}}(s,x)$. Note that this inequality is only useful for $\gamma > 0$, in which case $\mathbb{X}_\theta^\star(s)$ is in fact a singleton. It is straightforward to verify that all loss functions (4) are non-negative and evaluate to zero if and only if $x \in \mathbb{X}_\theta^\star(s)$. In the case of the first-order loss, for instance, this equivalence holds because the first-order condition (7) is both necessary and sufficient for the optimality of x. We remark that (6) remains valid if γ depends on s and θ. □

While the basic estimation models in statistical learning all minimize an empirical loss as in (2), some inverse optimization models proposed in [11] implicitly minimize the worst-case loss across all training samples. To capture both approaches in a unified model, we suggest here to minimize a normalized, positive homogeneous and monotone risk measure ρ that penalizes positive losses. More precisely, we denote by $\rho^{\mathbb{Q}}(\ell_\theta)$ the risk of the loss $\ell_\theta(\xi)$ if $\xi = (s,x)$ follows the distribution $\mathbb{Q}$. The inverse optimization problem (2) thus generalizes to

$$\underset{\theta\in\Theta}{\text{minimize}}\;\; \rho^{\widehat{\mathbb{P}}_N}(\ell_\theta), \qquad \text{where} \qquad \widehat{\mathbb{P}}_N := \frac{1}{N}\sum_{i=1}^N \delta_{\widehat{\xi}_i} \tag{8}$$

represents the *empirical distribution* on the training samples. In the remainder we refer to $\rho^{\widehat{\mathbb{P}}_N}(\ell_\theta)$ as the *empirical* or *in-sample risk*. Note that (8) reduces indeed to (2) if we choose the expected value as the risk measure. Two alternative risk measures that could be used in (8) are described in the following example.

Example 1 (Risk measures) A popular risk measure that the observer could use to quantify the risk of a positive loss is the *conditional value-at-risk* (CVaR) at level $\alpha \in (0, 1]$, which is defined as

$$\mathrm{CVaR}_\alpha^{\mathbb{Q}}(\ell_\theta) = \inf_\tau \ \tau + \frac{1}{\alpha}\mathbb{E}^{\mathbb{Q}}\big[\max\{\ell_\theta(s,x) - \tau, 0\}\big], \tag{9a}$$

see [33]. For $\alpha = 1$, the CVaR reduces to the expected value, and for $\alpha \downarrow 0$, it converges to the essential supremum of the loss. Alternatively, the observer could use the *value-at-risk* (VaR) at level $\alpha \in [0, 1]$ defined as

$$\mathrm{VaR}_\alpha^{\mathbb{Q}}(\ell_\theta) = \inf_\tau \ \big\{\tau \ : \ \mathbb{Q}\big[\ell_\theta(s,x) \le \tau\big] \ge 1 - \alpha\big\}. \tag{9b}$$

Note that the VaR coincides with the upper $(1 - \alpha)$-quantile of the loss distribution. Moreover, if $\ell_\theta(s, x)$ has a continuous marginal distribution under $\mathbb{Q}$, then $\mathrm{CVaR}_\alpha^{\mathbb{Q}}(\ell_\theta)$ coincides with the expected loss above $\mathrm{VaR}_\alpha^{\mathbb{Q}}(\ell_\theta)$.

By definition, the loss function $\ell_\theta(\xi)$ is non-negative for all $\xi \in \Xi$ and $\theta \in \Theta$. The monotonicity and normalization of the risk measure ρ thus imply that

$$\rho^{\widehat{\mathbb{P}}_N}(\ell_\theta) \ge \rho^{\widehat{\mathbb{P}}_N}(0) = 0 \qquad \forall \theta \in \Theta,$$

which in turn implies that the optimal value of problem (8) is necessarily larger than or equal to zero. Moreover, if the agent's true objective function is contained in $\mathcal{F}$, that is, if $F = F_{\theta^\star}$ for some $\theta^\star \in \Theta$, then the loss $\ell_{\theta^\star}(\widehat{\xi}_i)$ vanishes for all i, indicating that the optimal value of (8) is zero and that $\theta^\star$ is optimal in (8). In this case, the optimal *value* of the in-sample risk minimization problem (8) is known *a priori*, and the only informative output of any solution scheme is an *optimizer*, that is, a model $\theta' \in \Theta$ with zero in-sample risk. If the number of training samples is moderate, then there may be multiple optimal solutions, and θ' may differ from the agent's true model $\theta^\star$.

Remark 2.6 (Choice of risk measures) If $F = F_{\theta^\star}$ for $\theta^\star \in \Theta$, then the minimum of (8) vanishes and is minimized by $\theta^\star$ irrespective of ρ. Thus, one might believe that the choice of the risk measure is immaterial for the inverse optimization problem. However, different risk measures may result in different solution sets. For example, if ρ is the CVaR at level $\alpha \in (0, 1]$, then θ' is a minimizer of (8) if and only if $\ell_{\theta'}(\widehat{s}_i, \widehat{x}_i) = 0$ for all $i \le N$. In contrast, if ρ is the VaR at level $\alpha \in [0, 1]$, then θ' is a minimizer of (8) if and only if $\ell_{\theta'}(\widehat{s}_i, \widehat{x}_i) = 0$ for a portion of at least $1 - \alpha$ of all N training samples. Thus, the use of VaR may lead to an inflated solution set. Moreover, the choice of ρ impacts the tractability of (8); see Sects. 5 and 6.

3 Inverse optimization under imperfect information

The proposed framework for inverse optimization described in Sect. 2 is predicated on the assumption of perfect information. Specifically, it is assumed that the agent is able to determine and implement the best response x to any given signal s and

that the observer can measure s and x precisely. Moreover, it is implicitly assumed that the family $\mathcal{F}$ of candidate objective functions contains the agent's true objective function F. In practice, however, the observer may be confronted with the following challenges:

(i) *Model uncertainty* The hypothesis space $\mathcal{F}$ chosen by the observer may not be rich enough to contain the agent's true objective function F or any strictly increasing transformation of F that encodes the same preferences.
(ii) *Measurement noise* The observed signal-response pairs may be corrupted by noise, which prevents the observer from measuring them exactly.
(iii) *Bounded rationality* The agent may settle for a suboptimal decision x due to cognitive or computational limitations. Even if the best response can be computed exactly, an exact implementation of the desired best response may not be possible due to implementation errors [9].

In the presence of *model uncertainty*, that is, if neither F nor any strictly increasing transformation of F is contained in the chosen hypothesis space $\mathcal{F}$, then there exists typically no $\theta \in \Theta$ such that the loss functions described in Sect. 2 vanish on all training samples. In this case a perfect recovery of the agent's preferences is fundamentally impossible, and the optimal value of the empirical risk minimization problem (8) is positive. The best the observer can hope for is to learn the parametric hypothesis that most accurately (but imperfectly) captures the agent's true preferences.

In the presence of *measurement noise* the observed training samples $\widehat{\xi}_i = (\widehat{s}_i, \widehat{x}_i)$ represent random perturbations of some unobservable pure samples $\widetilde{\xi}_i = (\widetilde{s}_i, \widetilde{x}_i)$. While $\widetilde{x}_i$ is an exact optimal response to an unperturbed signal $\widetilde{s}_i$, the noisy response $\widehat{x}_i$ generically constitutes a suboptimal—maybe even infeasible—response to $\widehat{s}_i$. We will henceforth assume that the noisy samples $\widehat{\xi}_i$ are mutually independent and follow an *in-sample distribution* $\mathbb{P}_{\text{in}}$, while the corresponding unperturbed samples $\widetilde{\xi}_i$ are governed by an *out-of-sample distribution* $\mathbb{P}_{\text{out}}$. While the distribution $\mathbb{P}_{\text{out}}$ of the perfect samples is supported on Ξ by construction, the in-sample distribution $\mathbb{P}_{\text{in}}$ of the noisy samples may or may not be supported on Ξ. If the noisy samples can materialize outside of Ξ, then we call the noise *inconsistent*. If all noisy samples are guaranteed to reside within Ξ, on the other hand, then we call the noise *consistent*. Thus, in the presence of measurement noise, the observer faces the challenging task to learn $\mathbb{P}_{\text{out}}$ from samples of $\mathbb{P}_{\text{in}}$. Note that in the absence of noise we have $\mathbb{P}_{\text{in}} = \mathbb{P}_{\text{out}}$, which coincides with the distribution $\mathbb{P}$ introduced in Sect. 2.

Agents suffering from *bounded rationality* may not be able to solve (1) to global optimality. Such agents may respond to a given signal s with a δ-suboptimal solution x_δ, that is, a decision $x_\delta \in \mathbb{X}(s)$ with $F(s, x_\delta) \leq \min_{x \in \mathbb{X}(s)} F(s, x) + \delta$. Here, the parameter $\delta \geq 0$ quantifies the agent's irrationality. Indeed, if $\delta > 0$, the agent accepts a cost increase of up to δ for the freedom of choosing a δ-suboptimal decision, which may require fewer cognitive or computational resources than finding a global minimizer. Note that the observer may mistakenly perceive the effects of bounded rationality as (consistent) measurement noise or vice versa. So one might argue that there is no need to distinguish between these two types of imperfect information. However, bounded rationality is fundamentally different from measurement noise if the observer *knows* that the imperfect measurements are caused by bounded rational-

ity. If the imperfections originate from measurement noise, then the observer aims to filter out the noise in order to predict the agent's pure decisions. If the imperfections originate from bounded rationality, on the other hand, then the observer aims to predict the agent's δ-suboptimal decisions and not the ideal global optimizers. In other words, the observer always aims to predict the agent's responses bar measurement noise, irrespective of whether these responses are rational or not. An observer who knows the agent's degree of irrationality δ may thus improve the predictive power of her learning model by replacing the suboptimality loss (4b) with the bounded rationality loss.

Definition 3.1 (*Bounded rationality loss*) The bounded rationality loss of model θ is given by

$$\ell_\theta(\delta; s, x) := \max\left\{F_\theta(s, x) - \min_{y \in \mathbb{X}(s)} F_\theta(s, y) - \delta, 0\right\}. \tag{10}$$

It quantifies the amount by which the suboptimality of x with respect to F_θ given signal s exceeds $\delta \geq 0$. By construction, $\ell_\theta(\delta; s, x) = 0$ whenever x is a δ-supoptimal response to the signal s, and $\ell_\theta(s, x) > 0$ otherwise.

Remark 3.2 (Estimating δ) Note that the bounded rationality loss with $\delta = 0$ reduces to the usual suboptimality loss (4b). Even if it is known that the agent suffers from bounded rationality, the constant δ is unlikely to be available in practice. However, as long as there are no other sources of imperfect information (such as model uncertainty or measurement noise), the observer can estimate δ and θ from the training data by solving

$$\underset{\theta \in \Theta,\, \delta \geq 0}{\text{minimize}} \{\delta : \ell_\theta(\delta; \widehat{s}_i, \widehat{x}_i) = 0 \ \forall i \leq N\}. \tag{11}$$

This variant of the inverse optimization problem identifies the smallest bounded rationality constant δ that explains all observed responses as δ-suboptimal decisions under model θ. Note that (11) constitutes a convex optimization problem as long as Θ is convex and the hypotheses $F_\theta(s, x)$ are affinely parameterized in θ, which guarantees that $\ell_\theta(\delta; s, x)$ is jointly convex in θ and δ for any fixed s and x. Moreover, instead of solving (11), one could equivalently solve the empirical risk minimization problem (8) with the suboptimality loss and the essential supremum risk measure to find θ and then set δ to the resulting optimal objective value.

To some extent, all of the complications discussed in this section are inevitable in any real application. In fact, they are likely to reflect the norm rather than the exception. Thus, the main objective of this paper is to develop inverse optimization models that can cope with imperfect information in a principled manner.

4 Distributionally robust inverse optimization

By combining different loss functions with different risk measures, one may synthesize different empirical risk minimization problems of the form (8). By construction, any solution of (8) has minimum in-sample risk. However, the in-sample risk merely

captures historical performance and is therefore of little practical interest. Instead, the observer seeks models that display promising performance on unseen future data.

Definition 4.1 (*Out-of-sample risk*) We refer to $\rho^{\mathbb{P}_{\text{out}}}(\ell_\theta)$ as the out-of-sample risk of the model $\theta \in \Theta$.

Ideally, the observer would want to minimize the out-of-sample risk over all candidate models $\theta \in \Theta$. This is impossible, however, because the out-of-sample distribution $\mathbb{P}_{\text{out}}$ of the signal-response pairs is unknown, and only a finite set of training samples from $\mathbb{P}_{\text{in}}$ is available (recall that $\mathbb{P}_{\text{in}} = \mathbb{P}_{\text{out}}$ if measurements are perfect). In this situation, the observer has to settle for a data-driven solution $\widehat{\theta}_N \in \Theta$ that depends on the training samples and attains—hopefully—a low out-of-sample risk. We emphasize that, due to its dependence on the training samples, $\widehat{\theta}_N$ constitutes a random object, whose stochastics is governed by the product distribution $\mathbb{P}_{\text{in}}^N$. A simple data-driven solution is obtained, for instance, by solving the empirical risk minimization problem (8).

While it is impossible to minimize the out-of-sample risk on the basis of the training samples, it is sometimes possible to establish data-driven out-of-sample guarantees in the sense of the following definition.

Definition 4.2 (*Out-of-sample guarantee*) We say that a data-driven solution $\widehat{\theta}_N$ enjoys an out-of-sample performance guarantee at significance level $\beta \in [0, 1]$ if there exists a data-driven certificate $\widehat{J}_N$ with

$$\mathbb{P}_{\text{in}}^N\big[\rho^{\mathbb{P}_{\text{out}}}(\ell_{\widehat{\theta}_N}) \leq \widehat{J}_N\big] \geq 1 - \beta. \tag{12}$$

Note that the probability in (12) is evaluated with respect to the distribution of N independent (potentially noisy) training samples, which impact both the data-driven solution $\widehat{\theta}_N$ and the certificate $\widehat{J}_N$. Note also that the certificate $\widehat{J}_N$ can be viewed as an upper $(1-\beta)$-confidence bound on the out-of-sample risk of $\widehat{\theta}_N$. Thus, we sometimes refer to the confidence level $1 - \beta$ as the certificate's *reliability*.

As the ideal goal to minimize the out-of-sample risk is unachievable, the observer might settle for the more modest goal to determine a data-driven solution that admits a low certificate with a high reliablity. We will now argue that this secondary goal is achievable by adopting a distributionally robust approach. Specifically, we will use the N training samples to design an ambiguity set $\widehat{\mathcal{P}}_N$ that contains the out-of-sample distribution $\mathbb{P}_{\text{out}}$ of the (perfect) signal-response pairs with confidence $1 - \beta$. Next, we construct the data-driven solution $\widehat{\theta}_N$ and the corresponding certificate $\widehat{J}_N$ by minimizing the worst-case risk across all models $\theta \in \Theta$, where the worst case is taken with respect to all signal-response distributions in the ambiguity set $\widehat{\mathcal{P}}_N$, that is, we set

$$\widehat{\theta}_N \in \arg\min_{\theta\in\Theta} \sup_{\mathbb{Q}\in\widehat{\mathcal{P}}_N} \rho^{\mathbb{Q}}(\ell_\theta) \quad \text{and} \quad \widehat{J}_N := \min_{\theta\in\Theta} \sup_{\mathbb{Q}\in\widehat{\mathcal{P}}_N} \rho^{\mathbb{Q}}(\ell_\theta). \tag{13}$$

It is clear that if $\mathbb{P}_{\text{in}}^N[\mathbb{P}_{\text{out}} \in \widehat{\mathcal{P}}_N] \geq 1 - \beta$, then the distributionally robust solution $\widehat{\theta}_N$ and the corresponding certificate $\widehat{J}_N$ defined above satisfy the out-of-sample guarantee (12). In order to ensure that $\widehat{\mathcal{P}}_N$ contains the unknown out-of-sample distribution

$\mathbb{P}_{\text{out}}$ with confidence $1-\beta$, we construct the ambiguity set $\widehat{\mathcal{P}}_N$ as a ball in the space of probability distributions with respect to the Wasserstein metric as suggested in [30].

Definition 4.3 (*Wasserstein metric*) For any integer $p \geq 1$ and closed set $\Xi \subset \mathbb{R}^{n+m}$ we let $\mathcal{M}^p(\Xi)$ be the space of all probability distributions $\mathbb{Q}$ supported on Ξ with $\mathbb{E}^{\mathbb{Q}}\big[\|\xi\|^p\big] = \int_\Xi \|\xi\|^p \, \mathbb{Q}(\mathrm{d}\xi) < \infty$. The p-Wasserstein distance between two distributions $\mathbb{Q}_1, \mathbb{Q}_2 \in \mathcal{M}^p(\mathbb{R}^{n+m})$ is defined as

$$\mathrm{W}_p(\mathbb{Q}_1, \mathbb{Q}_2) := \inf \left\{ \left(\int \|\xi_1 - \xi_2\|^p \, \Pi(\mathrm{d}\xi_1, \mathrm{d}\xi_2) \right)^{1/p} : \begin{array}{l} \Pi \text{ is a joint distribution of } \xi_1 \text{ and } \xi_2 \\ \text{with marginals } \mathbb{Q}_1 \text{ and } \mathbb{Q}_2, \text{ respectively} \end{array} \right\}.$$

The Wasserstein distance $\mathrm{W}_p\big(\mathbb{Q}_1, \mathbb{Q}_2\big)$ can be viewed as the (p-th root of the) minimum cost for moving the distribution $\mathbb{Q}_1$ to $\mathbb{Q}_2$, where the cost of moving a unit mass from ξ_1 to ξ_2 amounts to $\|\xi_1 - \xi_2\|^p$. The joint distribution Π of ξ_1 and ξ_2 is therefore naturally interpreted as a mass transportation plan.

We define the ambiguity set as a p-Wasserstein ball in $\mathcal{M}^p(\Xi)$ centered at the empirical distribution $\widehat{\mathbb{P}}_N$ defined in (8). Specifically, we define the p-Wasserstein ball of radius ε around $\widehat{\mathbb{P}}_N$ as

$$\mathbb{B}_\varepsilon^p(\widehat{\mathbb{P}}_N) := \big\{ \mathbb{Q} \in \mathcal{M}^p(\Xi) \; : \; \mathrm{W}_p\big(\mathbb{Q}, \widehat{\mathbb{P}}_N\big) \leq \varepsilon \big\}.$$

Note that if $\varepsilon = 0$ and the empirical distribution is supported on Ξ, which is necessarily true in the absence of measurement noise, then the Wasserstein ball $\mathbb{B}_\varepsilon^p(\widehat{\mathbb{P}}_N)$ shrinks to the singleton set that contains only the empirical distribution. In this case, the distributionally robust inverse optimization problem (13) reduces to the empirical risk minimization problem (8). In order to establish out-of-sample guarantees, we must assume that the p-Wasserstein distance between $\mathbb{P}_{\text{in}}$ and $\mathbb{P}_{\text{out}}$ is bounded by a known constant $\varepsilon_0 \geq 0$. This is the case, for instance, if the noise is additive and all noise realizations are bounded by ε_0 with certainty.

Proposition 4.4 (Measure concentration) *Assume that there exists $a > 1$ with $A := \mathbb{E}^{\mathbb{P}_{\text{in}}}[\exp(\|\xi\|^a)] < \infty$. Assume also that $\beta \in (0,1)$ is a prescribed significance level, $\mathrm{W}_p\big(\mathbb{P}_{\text{in}}, \mathbb{P}_{\text{out}}\big) \leq \varepsilon_0$, and $m + n \neq 2p$.*[1] *Then, there exist constants $c_1, c_2 > 0$ that depend only on a, A, m and n such that $\mathbb{P}_{\text{in}}^N[\mathbb{P}_{\text{out}} \in \mathbb{B}_\varepsilon^p(\widehat{\mathbb{P}}_N)] \geq 1 - \beta$ whenever $\varepsilon \geq \varepsilon_0 + \varepsilon_N(\beta)$, where*

$$\varepsilon_N(\beta) := \begin{cases} \left(\frac{\log(c_1 \beta^{-1})}{c_2 N}\right)^{\min\left\{p(m+n)^{-1}, \frac{1}{2}\right\}} & \text{if } N \geq \frac{\log(c_1\beta^{-1})}{c_2}, \\ \left(\frac{\log(c_1 \beta^{-1})}{c_2 N}\right) & \text{if } N < \frac{\log(c_1\beta^{-1})}{c_2}. \end{cases} \tag{14}$$

[1] Proposition 4.4 readily extends to the case $n+m = 2p$ at the expense of additional notation by leveraging [21, Theorem 2].

Proof Select any $\varepsilon \geq \varepsilon_0 + \varepsilon_N(\beta)$. Then, the triangle inequality implies

$$\mathrm{W}_p\big(\mathbb{P}_{\mathrm{out}}, \widehat{\mathbb{P}}_N\big) \leq \mathrm{W}_p\big(\mathbb{P}_{\mathrm{out}}, \mathbb{P}_{\mathrm{in}}\big) + \mathrm{W}_p\big(\mathbb{P}_{\mathrm{in}}, \widehat{\mathbb{P}}_N\big).$$

By assumption, the first term on the right hand side is bounded by ε_0 with certainty. Theorem 3.5 in [30], which leverages a powerful measure concentration result developed in [21, Theorem 2], further guarantees that the second term is bounded above by $\varepsilon_N(\beta)$ with confidence $1-\beta$. As $\mathbb{P}_{\mathrm{out}}$ is supported on Ξ, we may thus conclude that $\mathbb{P}_{\mathrm{out}} \in \mathbb{B}_\varepsilon^p(\widehat{\mathbb{P}}_N)$ with probability $1-\beta$. This observation completes the proof. □

Proposition 4.4 ensures that the distributionally robust solution $\widehat{\theta}_N$ and the corresponding certificate $\widehat{J}_N$ induced by a Wasserstein ambiguity set of radius $\varepsilon \geq \varepsilon_0 + \varepsilon_N(\beta)$ satisfy the out-of-sample guarantee (12). One can further show that if there is no measurement noise ($\mathbb{P}_{\mathrm{in}} = \mathbb{P}_{\mathrm{out}} = \mathbb{P}$) while $\beta_N \in (0,1)$ for $N \in \mathbb{N}$ satisfies $\sum_{N=1}^{\infty} \beta_N < \infty$ and $\lim_{N\to\infty} \varepsilon_N(\beta_N) = 0$,[2] then any accumulation point of $\{\widehat{\theta}_N\}_{N\in\mathbb{N}}$ is $\mathbb{P}^\infty$-almost surely a minimizer of the out-of-sample risk $\rho^{\mathbb{P}}(\ell_\theta)$ over $\theta \in \Theta$; see [30, Theorem 3.6].

We emphasize that asymptotic consistency is lost when $\mathbb{P}_{\mathrm{in}}$ differs from $\mathbb{P}_{\mathrm{out}}$. In this case it is fundamentally impossible to systematically recover a minimizer of the out-of-sample risk $\rho^{\mathbb{P}}(\ell_\theta)$ even if infinitely many samples from $\mathbb{P}_{\mathrm{in}}$ are available. However, it is possible to construct examples where the solution of the distributionally robust inverse optimization problem (13) strictly outperforms that of the empirical risk minimization model (8) in out-of-sample tests even if the number of training samples tends to infinity.

Besides offering rigorous out-of-sample and asymptotic guarantees, the proposed approach to distributionally robust inverse optimization can be shown to be tractable if the search space contains only linear or quadratic hypotheses, and risk is measured by the CVaR of the suboptimality loss (4b) or the first-order loss (4c). Unfortunately, tractability is lost when minimizing the predictability loss (4a), which is the actual quantitiy of interest. However, the computable distributionally robust solutions $(\widehat{\theta}_N^{\mathrm{s}}, \widehat{J}_N^{\mathrm{s}})$ and $(\widehat{\theta}_N^{\mathrm{f}}, \widehat{J}_N^{\mathrm{f}})$ corresponding to the suboptimality and first-order losses, respectively, can be used to construct out-of-sample guarantees for the predictability loss if the hypotheses are uniformly strongly convex with parameter $\gamma > 0$. Indeed, if ℓ_θ^{p}, ℓ_θ^{s} and ℓ_θ^{f} denote the predictability, suboptimality and first-order losses, respectively, we have

$$\mathbb{P}_{\mathrm{in}}^N\left[\rho^{\mathbb{P}_{\mathrm{out}}}(\ell^{\mathrm{s}}_{\widehat{\theta}^{\mathrm{s}}_N}) \leq \widehat{J}^{\mathrm{s}}_N\right] \geq 1-\beta \quad \Longrightarrow \quad \mathbb{P}_{\mathrm{in}}^N\left[\rho^{\mathbb{P}_{\mathrm{out}}}(\ell^{\mathrm{p}}_{\widehat{\theta}^{\mathrm{s}}_N}) \leq \frac{2}{\gamma}\widehat{J}^{\mathrm{s}}_N\right] \geq 1-\beta \tag{15a}$$

and

$$\mathbb{P}_{\mathrm{in}}^N\left[\rho^{\mathbb{P}_{\mathrm{out}}}(\ell^{\mathrm{f}}_{\widehat{\theta}^{\mathrm{f}}_N}) \leq \widehat{J}^{\mathrm{f}}_N\right] \geq 1-\beta \quad \Longrightarrow \quad \mathbb{P}_{\mathrm{in}}^N\left[\rho^{\mathbb{P}_{\mathrm{out}}}(\ell^{\mathrm{p}}_{\widehat{\theta}^{\mathrm{f}}_N}) \leq \frac{2}{\gamma}\widehat{J}^{\mathrm{f}}_N\right] \geq 1-\beta, \tag{15b}$$

[2] A possible choice is $\beta_N = \exp(-\sqrt{N})$.

where both implications follow from the dominance relation (6) established in Proposition 2.5 and the scale invariance and monotonicity of the CVaR. Thus, both $\widehat{\theta}_N^{\mathrm{s}}$ and $\widehat{\theta}_N^{\mathrm{f}}$ are efficiently computable and offer an out-of-sample guarantee for the predictability loss. While both guarantees involve the same confidence level $1-\beta$, however, the underlying certificates J_N^{s} and J_N^{f} are generically different. One can again use the dominance relation (6) from Proposition 2.5 and the monotonicity of the CVaR to show that $J_N^{\mathrm{s}} \leq J_N^{\mathrm{f}}$. Thus, minimizing the suboptimality loss results in a weakly stronger predictability guarantee. This reasoning suggests that the observer should favor the suboptimality loss (4b) over the first-order loss (4c). We emphasize that unlike the suboptimality loss, however, the first-order loss leads to tractable inverse optimization models even in the presence of *several* strategically interacting agents [11].

Remark 4.5 (Out-of-sample guarantees in [11]) The out-of-sample guarantees provided in [11] only apply to the VaR, and an extension to other risk measures is not envisaged. Specifically, [11, Theorem 6] provides an out-of-sample guarantee for the VaR of the first-order loss, while [11, Theorem 6] offers an out-of-sample guarantee for the VaR of the predictability loss. In contrast, the distributionally robust approach discussed here offers out-of-sample guarantees for any normalized, positive homogeneous and monotone risk measure including the VaR or any coherent risk measure such as the CVaR etc.

Remark 4.6 (Curse of dimensionality) Proposition 4.4 implies that the distributionally robust solution $\widehat{\theta}_N$ and its certificate $\widehat{J}_N$ satisfy the out-of-sample guarantee (12) for $\mathbb{P}_{\mathrm{in}} = \mathbb{P}_{\mathrm{out}}$ if the Wasserstein radius scales as $\varepsilon_N(\beta) \propto \mathcal{O}(N^{-1/(m+n)})$. This entails a curse of dimensionality, that is, both $\varepsilon_N(\beta)$ as well as $\widehat{J}_N$ converge slowly for large input and/or output dimensions. By generalizing [36, Theorem 4.6], however, one can show that a significantly smaller Wasserstein radius of the order $\mathcal{O}((m+n)\sqrt{\log(N)/N})$ is sufficient for (12) whenever Θ is compact and does not contain 0, while $\ell_\theta(\xi)$ is Lipschitz continuous in $\theta \in \Theta$ for every $\xi \in \Xi$.

5 Linear hypotheses

On the one hand, the hypothesis space $\mathcal{F}$ should be rich enough to contain the agent's unknown true objective function F. On the other hand, $\mathcal{F}$ should be small enough to ensure tractability of the distributionally robust inverse optimization problem (13) and to prevent degeneracy of its optimal solutions. A particular class $\mathcal{F}$ that strikes this delicate balance and proves useful in many applications is the family of linear objective functions $F_\theta(s,x) := \langle \theta, x \rangle$. The corresponding search space $\Theta \subseteq \mathbb{R}^n$ may account for prior information on the agent's objective and should not contain $\theta = 0$, which corresponds to a trivial constant objective function that renders every response optimal. Examples of tractable search spaces are listed below.

Example 2 (Tractable search spaces) If there is no prior information on F, it is natural to set

$$\Theta := \big\{\theta \in \mathbb{R}^n : \|\theta\|_\infty = 1\big\}. \tag{16a}$$

Note that the normalization $\|\theta\|_\infty = 1$ is non-restrictive because the objective functions corresponding to θ and $\kappa\,\theta$ imply the same preferences for any model $\theta \neq 0$ and scaling factor $\kappa > 0$. In fact, we could define Θ as the unit sphere induced by any norm on $\mathbb{R}^n$. However, the ∞-norm stands out from a computational perspective. While all norm spheres are non-convex and therefore a priori unattractive as search spaces, the ∞-norm sphere decomposes into $2n$ polytopes—one for each facet. This polyhedral decomposition property allows us to optimize efficiently over Θ.

If F is known to be non-decreasing in the agent's decisions, a natural choice is

$$\Theta := \{\theta \in \mathbb{R}^n : \|\theta\|_1 = 1, \theta \geq 0\}. \tag{16b}$$

This search space has been used in [27] and constitutes a single convex polytope.

If F is believed to reside in the vicinity of a nominal objective function $\langle \theta_0, x \rangle$ as in [4], then we may set

$$\Theta := \big\{\theta \in \mathbb{R}^n : \|\theta - \theta_0\| \leq \Gamma\big\}, \tag{16c}$$

where $\|\cdot\|$ denotes a generic norm, and Γ reflects the degree of uncertainty about the nominal model θ_0.

When focusing on linear hypotheses, the suboptimality loss function (4b) reduces to

$$\begin{aligned} F_\theta(s,x) - \min_{y\in\mathbb{X}(s)} F_\theta(s,y) &= \langle \theta, x\rangle - \min_{y\in\mathbb{X}(s)} \langle \theta, y\rangle = \max_{y\in\mathbb{X}(s)} \langle \theta, x-y\rangle \\ &= \max_{y\in\mathbb{X}(s)} \langle \nabla_x F_\theta(s,x), x-y\rangle \end{aligned} \tag{17}$$

and thus equals the first-order loss (4c), which is positive homogeneous and subadditive in θ. The tractability results to be established below rely on the following assumption.

Assumption 5.1 (*Conic representable support*) The signal space $\mathbb{S}$ and the feasible set $\mathbb{X}(s)$ are conic representable, that is,

$$\mathbb{S} = \big\{s \in \mathbb{R}^m : Cs \succeq_{\mathcal{C}} d\big\} \quad \text{and} \quad \mathbb{X}(s) = \{x \in \mathbb{R}^n : Wx \succeq_{\mathcal{K}} Hs + h\} \quad \forall s \in \mathbb{S},$$

where the relations '$\succeq_{\mathcal{C}}$' and '$\succeq_{\mathcal{K}}$' represent conic inequalities with respect to some proper convex cones $\mathcal{C}$ and $\mathcal{K}$ of appropriate dimensions, respectively. The set Ξ of all possible signal-response pairs thus reduces to

$$\Xi = \big\{(s,x) \in \mathbb{R}^m \times \mathbb{R}^n : Cs \succeq_{\mathcal{C}} d,\ Wx \succeq_{\mathcal{K}} Hs + h\big\}.$$

We also assume that the convex set Ξ admits a Slater point.

Under Assumption 5.1, the suboptimality loss $\ell_\theta(s,x)$ is concave in (s,x) for every fixed θ, see, e.g., [14, Section 3.2.5]. We are now ready to state our first tractability result for the class of linear hypotheses.

Theorem 5.2 (Linear hypotheses and suboptimality loss) *Assume that $\mathcal{F}$ represents the class of linear hypotheses with a search space of the form (16) and that Assumption 5.1 holds. If the observer uses the suboptimality loss (4b) and measures risk using the* CVaR *at level $\alpha \in (0, 1]$, then the distributionally robust inverse optimization problem (13) over the 1-Wasserstein ball is equivalent to the finite conic program*[3]

$$
\begin{array}{ll}
\text{minimize} & \tau + \dfrac{1}{\alpha}\left(\varepsilon\lambda + \dfrac{1}{N}\displaystyle\sum_{i=1}^{N} r_i\right) \\
\text{subject to} & \theta \in \Theta,\quad \lambda \in \mathbb{R}_+,\quad \tau, r_i \in \mathbb{R},\quad \phi_{i1}, \phi_{i2} \in \mathcal{C}^*,\quad \mu_{i1}, \mu_{i2}, \gamma_i \in \mathcal{K}^* \qquad \forall i \leq N \\
& \left\langle C\widehat{s}_i - d, \phi_{i1}\right\rangle + \left\langle W\widehat{x}_i - H\widehat{s}_i - h, \mu_{i1} + \gamma_i\right\rangle \leq r_i + \tau \qquad \forall i \leq N \\
& \left\langle C\widehat{s}_i - d, \phi_{i2}\right\rangle + \left\langle W\widehat{x}_i - H\widehat{s}_i - h, \mu_{i2}\right\rangle \leq r_i,\quad \theta = W^\intercal\gamma_i \qquad \forall i \leq N \\
& \left\| \begin{pmatrix} C^\intercal\phi_{i1} - H^\intercal(\mu_{i1} + \gamma_i) \\ W^\intercal(\mu_{i1} + \gamma_i) \end{pmatrix} \right\|_* \leq \lambda,\quad \left\| \begin{pmatrix} C^\intercal\phi_{i2} - H^\intercal\mu_{i2} \\ W^\intercal\mu_{i2} \end{pmatrix} \right\|_* \leq \lambda \;\; \forall i \leq N.
\end{array}
\tag{18}
$$

We emphasize that Theorem 5.2 remains valid if the training samples are inconsistent with the given support information, that is, if $(\widehat{s}_i, \widehat{x}_i) \notin \Xi$ for some $i \leq N$, in which case $\ell_\theta(\widehat{s}_i, \widehat{x}_i)$ can even be negative.

Proof of Theorem 5.2 By the definition of CVaR, the objective function of (13) can be expressed as

$$
\begin{aligned}
\sup_{\mathbb{Q} \in \mathbb{B}_\varepsilon^1(\widehat{\mathbb{P}}_N)} \rho^{\mathbb{Q}}(\ell_\theta) &= \sup_{\mathbb{Q} \in \mathbb{B}_\varepsilon^1(\widehat{\mathbb{P}}_N)} \inf_{\tau} \tau + \frac{1}{\alpha}\mathbb{E}^{\mathbb{Q}}\big[\max\{\ell_\theta(s, x) - \tau, 0\}\big] \\
&= \inf_{\tau} \tau + \frac{1}{\alpha} \sup_{\mathbb{Q} \in \mathbb{B}_\varepsilon^1(\widehat{\mathbb{P}}_N)} \mathbb{E}^{\mathbb{Q}}\big[\max\{\ell_\theta(s, x) - \tau, 0\}\big].
\end{aligned}
\tag{19}
$$

The interchange of the maximization over $\mathbb{Q}$ and the minimization over τ in the second line is justified by Sion's minimax theorem [38], which applies because the Wasserstein ball $\mathbb{B}_\varepsilon^1(\widehat{\mathbb{P}}_N)$ is weakly compact [12, p. 2298]. The subordinate worst-case expectation problem in the second line of (19) constitutes a semi-infinite linear program. As the corresponding integrand is given by the maximum of $\ell_\theta(s, x) - \tau$ and 0, both of which can be viewed as proper concave functions in (s, x), this worst-case expectation problem admits a strong dual semi-infinite linear program of the form

$$
\begin{array}{lll}
\inf\limits_{\lambda \geq 0, r_i} & \varepsilon\lambda + \frac{1}{N}\displaystyle\sum_{i=1}^{N} r_i & \\
\text{s.t.} & \sup\limits_{(s,x) \in \Xi} \sup\limits_{y \in \mathbb{X}(s)} \left\langle \theta, x - y\right\rangle - \tau - \lambda\|(s, x) - (\widehat{s}_i, \widehat{x}_i)\| \leq r_i \;\; \forall i \leq N & \\
& \sup\limits_{(s,x) \in \Xi} -\lambda\|(s, x) - (\widehat{s}_i, \widehat{x}_i)\| \leq r_i & \forall i \leq N,
\end{array}
\tag{20}
$$

[3] Strictly speaking, if Θ is an ∞-norm ball of the form (16a), then (18) can be viewed as a family of $2n$ finite conic programs because Θ is non-convex but decomposes into $2n$ convex polytopes.

see Theorem 4.2 in [30] for a detailed derivation of (20) for more general integrands. By the definitions of $\mathbb{X}(s)$ and Ξ put forth in Assumption 5.1, respectively, the i-th member of the first constraint group in (20) holds if and only if the optimal value of the conic program

$$\begin{array}{ll} \sup\limits_{s,x,y} & \langle \theta, x - y \rangle - \tau - \lambda \|(s,x) - (\widehat{s}_i, \widehat{x}_i)\| \\ \text{s.t.} & Cs \succeq_{\mathcal{C}} d, \;\; Wx \succeq_{\mathcal{K}} Hs + h, \;\; Wy \succeq_{\mathcal{K}} Hs + h \end{array}$$

is smaller or equal to r_i. The dual of this conic program is given by

$$\begin{array}{ll} \inf & \langle C\widehat{s}_i - d, \phi_{i1} \rangle + \langle W\widehat{x}_i - H\widehat{s}_i - h, \mu_{i1} + \gamma_i \rangle - \tau \\ \text{s.t.} & \phi_{i1} \in \mathcal{C}^*, \;\; \mu_{i1}, \gamma_i \in \mathcal{K}^* \\ & \left\| \begin{pmatrix} C^\intercal \phi_{i1} - H^\intercal(\mu_{i1} + \gamma_i) \\ W^\intercal(\mu_{i1} + \gamma_i) \end{pmatrix} \right\|_* \leq \lambda, \quad \theta = W^\intercal \gamma_i, \end{array}$$

and strong duality holds because the uncertainty set Ξ contains a Slater point due to Assumption 5.1. Thus, the i-th member of the first constraint group in (20) holds if and only if there exist $\phi_{i1} \in \mathcal{C}^*$ and $\mu_{i1}, \gamma_i \in \mathcal{K}^*$ such that $\theta = W^\intercal \gamma_i$,

$$\begin{array}{l} \langle C\widehat{s}_i - d, \phi_{i1} \rangle + \langle W\widehat{x}_i - H\widehat{s}_i - h, \mu_{i1} + \gamma_i \rangle \leq r_i \\ \quad + \tau \text{ and } \left\| \begin{pmatrix} C^\intercal \phi_{i1} - H^\intercal(\mu_{i1} + \gamma_i) \\ W^\intercal(\mu_{i1} + \gamma_i) \end{pmatrix} \right\|_* \leq \lambda. \end{array}$$

A similar reasoning shows that the i-th member of the second constraint group in (20) holds if and only if there exist $\phi_{i2} \in \mathcal{C}^*$ and $\mu_{i2} \in \mathcal{K}^*$ such that

$$\langle C\widehat{s}_i - d, \phi_{i2} \rangle + \langle W\widehat{x}_i - H\widehat{s}_i - h, \mu_{i2} \rangle \leq r_i \quad \text{and} \quad \left\| \begin{pmatrix} C^\intercal \phi_{i2} - H^\intercal \mu_{i2} \\ W^\intercal \mu_{i2} \end{pmatrix} \right\|_* \leq \lambda.$$

In summary, the worst-case expectation in the second line of (19) thus coincides with the optimal value of the finite conic program

$$\begin{array}{lll} \inf & \varepsilon\lambda + \frac{1}{N} \sum\limits_{i=1}^{N} r_i & \\ \text{s.t.} & \lambda \in \mathbb{R}_+, \;\; r_i \in \mathbb{R}, \;\; \phi_{i1}, \phi_{i2} \in \mathcal{C}^*, \;\; \mu_{i1}, \mu_{i2}, \gamma_i \in \mathcal{K}^* & \forall i \leq N \\ & \langle C\widehat{s}_i - d, \phi_{i1} \rangle + \langle W\widehat{x}_i - H\widehat{s}_i - h, \mu_{i1} + \gamma_i \rangle \leq r_i + \tau & \forall i \leq N \\ & \langle C\widehat{s}_i - d, \phi_{i2} \rangle + \langle W\widehat{x}_i - H\widehat{s}_i - h, \mu_{i2} \rangle \leq r_i, \;\; \theta = W^\intercal \gamma_i & \forall i \leq N \\ & \left\| \begin{pmatrix} C^\intercal \phi_{i1} - H^\intercal(\mu_{i1} + \gamma_i) \\ W^\intercal(\mu_{i1} + \gamma_i) \end{pmatrix} \right\|_* \leq \lambda, \;\; \left\| \begin{pmatrix} C^\intercal \phi_{i2} - H^\intercal \mu_{i2} \\ W^\intercal \mu_{i2} \end{pmatrix} \right\|_* \leq \lambda & \forall i \leq N. \end{array}$$

The claim then follows by substituting this conic program into (19). □

If $(\widehat{s}_i, \widehat{x}_i) \in \Xi$ for all $i \leq N$, then the conic program (18) simplifies. In this case, the maximization problems in the last constraint group of (20) all evaluate to zero,

which implies that $\langle C\widehat{s}_i - d, \phi_{i2}\rangle + \langle W\widehat{x}_i - H\widehat{s}_i - h, \mu_{i2}\rangle \leq r_i$ reduces to $r_i \geq 0$, while the decision variables ϕ_{i2} and μ_{i2} as well as the constraints

$$\left\| \begin{pmatrix} C^\intercal \phi_{i2} - H^\intercal \mu_{i2} \\ W^\intercal \mu_{i2} \end{pmatrix} \right\|_* \leq \lambda$$

can be omitted from (18) for all $i \leq N$.

For stress test experiments it is often desirable to know the extremal distribution that achieves the worst-case risk in (13). The following theorem shows that this extremal distribution can be constructed systematically for any fixed $\theta \in \Theta$ by solving a finite convex optimization problem akin to (18).

Theorem 5.3 (Worst-case distribution for linear hypotheses) *Under the assumptions of Theorem 5.2, the worst-case risk in (18) corresponding to a fixed $\theta \in \Theta$ coincides with the optimal value of a finite convex program,* i.e.*,*

$$\begin{aligned} \sup_{\mathbb{Q}\in\mathbb{B}_\varepsilon^p(\widehat{\mathbb{P}}_N)} \mathrm{CVaR}_\alpha^{\mathbb{Q}}(\ell_\theta) = \max\ & \frac{1}{\alpha N}\sum_{i=1}^N \pi_{i1}\ell_\theta\left(\frac{p_{i1}}{\pi_{i1}}, \frac{q_{i1}}{\pi_{i1}}\right) \\ \text{s.t.}\ & \pi_{ij} \in \mathbb{R}_+, \quad p_{ij} \in \mathbb{R}^m, \quad q_{ij} \in \mathbb{R}^n && \forall i \leq N,\ j \leq 2 \\ & \tfrac{p_{ij}}{\pi_{ij}} \in \mathbb{S}, \quad \tfrac{q_{ij}}{\pi_{ij}} \in \mathbb{X}(\tfrac{p_{ij}}{\pi_{ij}}) && \forall i \leq N,\ j \leq \\ & \pi_{i1} + \pi_{i2} = 1, \quad \frac{1}{N}\sum_{i=1}^N \pi_{i1} = \alpha && \forall i \leq N \\ & \frac{1}{N}\sum_{i=1}^N\sum_{j=1}^2 \pi_{ij}\left\|\begin{pmatrix} \frac{p_{ij}}{\pi_{ij}} - \widehat{s}_i \\ \frac{q_{ij}}{\pi_{ij}} - \widehat{x}_i \end{pmatrix}\right\| \leq \varepsilon. \end{aligned} \tag{21}$$

For any optimal solution $\{\pi_{ij}^\star, p_{ij}^\star, q_{ij}^\star\}$ of this convex program, the discrete distribution

$$\mathbb{Q}^\star := \frac{1}{N}\sum_{i=1}^N\sum_{j=1}^2 \pi_{ij}^\star \delta_{\xi_{ij}^\star} \quad \textit{with} \quad \xi_{ij}^\star := \left(\frac{p_{ij}^\star}{\pi_{ij}^\star}, \frac{q_{ij}^\star}{\pi_{ij}^\star}\right)^\intercal$$

belongs to the Wassertein ball $\mathbb{B}_\varepsilon^1(\widehat{\mathbb{P}}_N)$ and attains the supremum on the left hand side of (21).

Proof As the loss function (17) is proper and jointly concave in x and s, we can use a similar reasoning as in [30, Theorem 4.5] to show that the convex program on the right hand side of (21) coincides with the strong dual of (18) for any fixed $\theta \in \Theta$. If this convex program is solvable and $\{\pi_{ij}^\star, p_{ij}^\star, q_{ij}^\star\}$ is a maximizer, then $\mathbb{Q}^\star \in \mathbb{B}_\varepsilon^1(\widehat{\mathbb{P}}_N)$ due to [30, Corollary 4.7]. It remains to be shown that $\mathrm{CVaR}_\alpha^{\mathbb{Q}^\star}(\ell_\theta)$ is no smaller than (18). Indeed, by the definition of CVaR we have

$$\begin{aligned}\mathrm{CVaR}_\alpha^{\mathbb{Q}^\star}(\ell_\theta) &= \inf_{\tau\in\mathbb{R}}\ \tau + \frac{1}{\alpha N}\sum_{i=1}^{N}\sum_{j=1}^{2}\pi_{ij}^\star \max\left\{\ell_\theta\left(\frac{p_{ij}^\star}{\pi_{ij}^\star},\frac{q_{ij}^\star}{\pi_{ij}^\star}\right)-\tau,0\right\}\\ &= \sup_{0\le v_{ij}\le\pi_{ij}^\star}\left\{\frac{1}{\alpha N}\sum_{i=1}^{N}\sum_{j=1}^{2}v_{ij}\ell_\theta\left(\frac{p_{ij}^\star}{\pi_{ij}^\star},\frac{q_{ij}^\star}{\pi_{ij}^\star}\right)\ :\ \alpha=\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{2}v_{ij}\right\}\\ &\ge \frac{1}{\alpha N}\sum_{i=1}^{N}\pi_{i1}^\star\ell_\theta\left(\frac{p_{i1}^\star}{\pi_{i1}^\star},\frac{q_{i1}^\star}{\pi_{i1}^\star}\right) = \sup_{\mathbb{Q}\in\mathbb{B}_\varepsilon^p(\widehat{\mathbb{P}}_N)}\mathrm{CVaR}_\alpha^{\mathbb{Q}}(\ell_\theta).\end{aligned}$$

Here, the second equality follows from strong linear programming duality, while the inequality follows from the feasibility of the solution $v_{i1}=\pi_{i1}^\star$ and $v_{i2}=0$ for $i\le N$.

We close this section by generalizing Theorem 5.2 to the bounded rationality loss (10). The proof largely parallels that of Theorem 5.2 and is therefore omitted for brevity.

Corollary 5.4 (Linear hypotheses and bounded rationality loss) *Assume that $\mathcal{F}$ represents the class of linear hypotheses with search space of the form (16) and that Assumptions 5.1 holds. If the observer uses the bounded rationality loss (10) and measures risk using the* CVaR *at level $\alpha\in(0,1]$, then the inverse optimization problem (13) over the 1-Wasserstein ball is equivalent to a variant of the conic program (18) with an additional decision variable $\tau\in\mathbb{R}_+$ and where the first constraint group is replaced with*

$$\left\langle C\widehat{s}_i-d,\phi_{i1}\right\rangle+\left\langle W\widehat{x}_i-H\widehat{s}_i-h,\mu_{i1}+\gamma_i\right\rangle\le r_i+\tau+\delta\quad\forall i\le N.$$

6 Quadratic hypotheses

Optimization problems with quadratic objectives abound in control [5], statistics [22], finance [29] and many other application domains. Algorithms for inverse optimization that can learn quadratic objective functions from signal-response pairs are therefore of great practical interest. This motivates us to consider the class $\mathcal{F}$ of quadratic hypotheses of the form $F_\theta(s,x):=\langle x,Q_{xx}x\rangle+\langle x,Q_{xs}s\rangle+\langle q,x\rangle$, which are encoded by a parameter $\theta:=(Q_{xx},Q_{xs},q)$. The corresponding search space should account for prior information on the agent's objective and should exclude $\theta=0$. Examples of tractable search spaces are listed below.

Example 3 (Tractable search spaces) If F is only known to be strongly convex in x, it is natural to set

$$\Theta:=\left\{\theta=(Q_{xx},Q_{xs},q)\in\mathbb{R}^{n\times n}\times\mathbb{R}^{n\times m}\times\mathbb{R}^n : Q_{xx}\succeq\mathbb{I}\right\}. \tag{22a}$$

Note that the normalization $Q_{xx}\succeq\mathbb{I}$ is non-restrictive because a positive scaling of the objective function does not alter the agent's preferences.

If F is only known to be bilinear in s and x, it is natural to set

$$\Theta := \left\{ \theta = (Q_{xx}, Q_{xs}, q) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^n : Q_{xx} \succeq 0,\ Q_{xs} = \mathbb{I} \right\}, \quad (22b)$$

where the normalization $Q_{xs} = \mathbb{I}$ can always be enforced by redefining s if necessary.

If F is close to a nominal objective function $\langle x, Q^0_{xx} x\rangle + \langle x, Q^0_{xs} s\rangle + \langle q^0, x\rangle$, then we may set

$$\Theta := \left\{ \theta = (Q_{xx}, Q_{xs}, q) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^n : Q_{xx} \succeq 0,\ \|\theta - \theta_0\| \le \Gamma \right\}, \quad (22c)$$

where $\|\cdot\|$ denotes a generic norm, and Γ captures the uncertainty of the nominal model $\theta_0 = (Q^0_{xx}, Q^0_{xs}, q^0)$.

When focusing on quadratic candidate objective functions, the suboptimality loss (4b) reduces to

$$\begin{aligned} \ell_\theta(s, x) &= \langle x, Q_{xx}x + Q_{xs}s + q\rangle - \min_{y\in\mathbb{X}(s)} \langle y, Q_{xx}y + Q_{xs}s + q\rangle \\ &= \max_{y\in\mathbb{X}(s)} \langle x, Q_{xx}x + Q_{xs}s + q\rangle - \langle y, Q_{xx}y + Q_{xs}s + q\rangle. \end{aligned} \quad (23)$$

As in Sect. 5, we suppose that Assumption 5.1 holds. In this setting the agent's decision problem (1) constitutes a conic program and is therefore tractable for common choices of the cones $\mathcal{C}$ and $\mathcal{K}$. In contrast, the inverse optimization problem (13) is hard. In fact, it is already hard to evaluate the objective function of (13) for a fixed θ. As we work with quadratic objectives, throughout this section we use the 2-norm on the signal-response space and the 2-Wasserstein metric to measure distances of distributions.

Theorem 6.1 (Intractability of (13) for quadratic hypotheses) *Assume that $\mathcal{F}$ represents the class of quadratic hypotheses with search space (22a) and that Assumption 5.1 holds. If the observer uses the suboptimality loss (4b) and measures risk using the* CVaR *at level $\alpha \in (0, 1]$, then evaluating the objective function of (13) for a fixed $\theta \in \Theta$ is NP-hard even if $N = 1$ (there is only one data point), $\alpha = 1$ (the observer is risk-neutral), $\mathbb{S}$ is a singleton, $\mathbb{X}(s)$ is a polytope independent of s, and $Q_{xs} = 0$.*

Proof The proof relies on a reduction from the NP-hard quadratic maximization problem [31].

Quadratic Maximization
Instance. A positive definite matrix $Q = Q^\intercal \succeq \mathbb{I}$.
Goal. Evaluate $\max_{\|x\|_\infty \le 1} \langle x, Qx\rangle$.

Given an input $Q \succeq \mathbb{I}$ to the quadratic maximization problem, we construct an instance of the inverse optimization problem (13) with $N = 1, \alpha = 1$, and Wasserstein radius $\varepsilon = \sqrt{n}$, where

$$\widehat{s}_1 := 0, \quad \widehat{x}_1 := 0, \quad Q_{xx} := Q, \quad Q_{xs} := 0, \quad \mathbb{S} := \{0\},$$

$$\mathbb{X}(s) := \{x \in \mathbb{R}^n : \|x\|_\infty \leq 1\}.$$

Under this parametrization, the objective function of (13) reduces to

$$\begin{aligned}\sup_{\mathbb{Q}\in\mathbb{B}_\varepsilon^2(\widehat{\mathbb{P}}_N)} \rho^{\mathbb{Q}}(\ell_\theta) &= \sup_{\mathbb{Q}\in\mathbb{B}_\varepsilon^2(\widehat{\mathbb{P}}_N)} \mathbb{E}^{\mathbb{Q}}\Big[\max_{y\in\mathbb{X}(s)} \big\langle x, Q_{xx}x + Q_{xs}s + q\big\rangle \\ &\qquad -\big\langle y, Q_{xx}y + Q_{xs}s + q\big\rangle\Big] \\ &= \sup_{\mathbb{Q}\in\mathbb{B}_\varepsilon^2(\widehat{\mathbb{P}}_N)} \mathbb{E}^{\mathbb{Q}}\Big[\max_{y\in\mathbb{X}(s)} \big\langle x, Qx\big\rangle - \big\langle y, Qy\big\rangle\Big] \\ &\leq \sup_{s\in\mathbb{S}, x\in\mathbb{X}(s)} \max_{y\in\mathbb{X}(s)} \big\langle x, Qx\big\rangle - \big\langle y, Qy\big\rangle = \sup_{\|x\|_\infty\leq 1} \big\langle x, Qx\big\rangle,\end{aligned}$$

where the inequality in the third line follows from the inclusion $\mathbb{B}_\varepsilon^2(\widehat{\mathbb{P}}_N) \subseteq \mathcal{M}^2(\Xi)$, while the last equality holds because the innermost maximum is attained at $y = 0$. As $(s, x) \in \Xi$ if and only if $s = 0$ and $\|x\|_\infty \leq 1$, we conclude that $(s, x) \in \Xi$ implies $\|x\|_2 \leq \sqrt{n}$ and $\|(s, x)\|_2^2 \leq n = \varepsilon^2$. Moreover, as the empirical distribution $\widehat{\mathbb{P}}_N$ coincides with the Dirac point measure at 0, the Wasserstein ball $\mathbb{B}_\varepsilon^2(\widehat{\mathbb{P}}_N)$ thus contains all distributions supported on Ξ, implying that the inequality in the above expression is in fact an equality. Hence, evaluating the objective function of (13) is tantamount to solving the NP-hard quadratic maximization problem. This observation completes the proof. □

Corollary 6.2 (Intractability of (13) for bilinear hypotheses) *If all assumptions of Theorem 6.1 hold but $\mathcal{F}$ denotes the class of bilinear hypotheses with search space (22b), then evaluating the objective of (13) for a fixed $\theta \in \Theta$ is NP-hard even if $N = 1$, $\alpha = 1$, $\mathbb{S}$ is a singleton, $\mathbb{X}(s)$ is a polytope independent of s and $Q_{xx} = 0$.*

Proof The proof is similar to that of Theorem 6.1 and omitted for brevity. □

Corollary 6.2 asserts that the inverse optimization problem (13) is intractable even if we focus on linear candidate objectives that depend on the exogenous signal s. This finding contrasts with the tractability Theorem 5.2 for candidate objectives independent of s. The intractability results portrayed in Theorem 6.1 and Corollary 6.2 motivate us to devise a safe conic approximation for the inverse optimization problem (13) with quadratic candidate objective functions.

Theorem 6.3 (Quadratic hypotheses and suboptimality loss) *Assume that $\mathcal{F}$ represents the class of quadratic hypotheses with a search space of the form (22) and that Assumption 5.1 holds. If the observer uses the suboptimality loss (4b) and measures risk using the* CVaR *at level $\alpha \in (0, 1]$, then the following conic program provides a*

safe approximation for the distributionally robust inverse optimization problem (13) *over the* 2*-Wasserstein ball:*

$$
\begin{array}{lll}
\textit{minimize} & \tau + \dfrac{1}{\alpha}\left(\varepsilon^2\lambda + \dfrac{1}{N}\displaystyle\sum_{i=1}^{N} r_i\right) & \\
\text{subject to} & \theta = (Q_{xx}, Q_{xs}, q) \in \Theta,\ \lambda \in \mathbb{R}_+,\ \tau, r_i, \rho_{i1}, \rho_{i2} \in \mathbb{R} & \forall i \le N \\
& \phi_{i1}, \phi_{i2} \in \mathcal{C}^*,\ \mu_{i1}, \mu_{i2}, \gamma_i \in \mathcal{K}^*,\ \chi_{i1}, \chi_{i2} \in \mathbb{R}^m,\ \zeta_{i1}, \eta_{i1}, \zeta_{i2} \in \mathbb{R}^n & \forall i \le N \\
& \chi_{i1} = \frac{1}{2}(-C^\intercal\phi_{i1} + H^\intercal(\mu_{i1} + \gamma_i) - 2\lambda\widehat{s}_i) & \forall i \le N \\
& \zeta_{i1} = \frac{1}{2}(-q - W^\intercal\mu_{i1} - 2\lambda\widehat{x}_i), \quad \eta_{i1} = \frac{1}{2}(q - W^\intercal\gamma_i) & \forall i \le N \\
& \rho_{i1} = \tau + r_i + \lambda\left(\langle\widehat{x}_i, \widehat{x}_i\rangle + \langle\widehat{s}_i, \widehat{s}_i\rangle\right) + \langle d, \phi_{i1}\rangle + \langle h, \mu_{i1} + \gamma_i\rangle & \forall i \le N \\
& \chi_{i2} = \frac{1}{2}(-C^\intercal\phi_{i2} + H^\intercal\mu_{i2} - 2\lambda\widehat{s}_i) & \forall i \le N \\
& \zeta_{i2} = \frac{1}{2}(-W^\intercal\mu_{i2} - 2\lambda\widehat{x}_i) & \forall i \le N \\
& \rho_{i2} = r_i + \lambda\left(\langle\widehat{x}_i, \widehat{x}_i\rangle + \langle\widehat{s}_i, \widehat{s}_i\rangle\right) + \langle d, \phi_{i2}\rangle + \langle h, \mu_i\rangle & \forall i \le N \\
& \begin{bmatrix} \lambda\mathbb{I} & -\frac{1}{2}Q_{xs}^\intercal & \frac{1}{2}Q_{xs}^\intercal & \chi_{i1} \\ -\frac{1}{2}Q_{xs} & \lambda\mathbb{I} - Q_{xx} & 0 & \zeta_{i1} \\ \frac{1}{2}Q_{xs} & 0 & Q_{xx} & \eta_{i1} \\ \chi_{i1}^\intercal & \zeta_{i1}^\intercal & \eta_{i1}^\intercal & \rho_{i1} \end{bmatrix} \succeq 0, \quad \begin{bmatrix} \lambda\mathbb{I} & 0 & \chi_{i2} \\ 0 & \lambda\mathbb{I} & \zeta_{i2} \\ \chi_{i2}^\intercal & \zeta_{i2}^\intercal & \rho_{i2} \end{bmatrix} \succeq 0 & \forall i \le N.
\end{array}
\tag{24}
$$

Note that Theorem 6.3 remains valid if $(\widehat{s}_i, \widehat{x}_i) \notin \Xi$ for some $i \le N$.

Proof of Theorem 6.3 As in the proof of Theorem 5.2 one can show that the objective function of the inverse optimization problem (13) coincides with a variant of (19) that involves the 2-Wasserstein ball. In the remainder, we derive a safe conic approximation for the subordinate worst-case expectation problem

$$
\sup_{\mathbb{Q}\in\mathbb{B}_\varepsilon^2(\widehat{\mathbb{P}}_N)} \mathbb{E}^{\mathbb{Q}}\big[\max\{\ell_\theta(s,x) - \tau, 0\}\big]. \tag{25}
$$

Duality arguments borrowed from [30, Theorem 4.2] imply that the above infinite-dimensional linear program admits a strong dual of the form

$$
\begin{array}{lll}
\displaystyle\inf_{\lambda\ge 0, r_i} & \varepsilon^2\lambda + \frac{1}{N}\displaystyle\sum_{i=1}^{N} r_i & \\
\text{s.t.} & \displaystyle\sup_{s\in\mathbb{S},\ x,y\in\mathbb{X}(s)} \big\langle x,\, Q_{xx}x + Q_{xs}s + q\big\rangle - \big\langle y,\, Q_{xx}y + Q_{xs}s + q\big\rangle - \tau & \\
& \qquad -\lambda\|(s,x) - (\widehat{s}_i, \widehat{x}_i)\|_2^2 \le r_i & \forall i \le N \\
& \displaystyle\sup_{s\in\mathbb{S},\ x\in\mathbb{X}(s)} -\lambda\|(s,x) - (\widehat{s}_i, \widehat{x}_i)\|_2^2 \le r_i & \forall i \le N.
\end{array}
\tag{26}
$$

By using the definitions of $\mathbb{S}$ and $\mathbb{X}(s)$ put forth in Assumption 5.1, the i-th member of the first constraint group in (26) is satisfied if and only if the optimal value of the maximization problem

$$\begin{array}{ll}\sup\limits_{s,x,y} & \langle x, Q_{xx}x + Q_{xs}s + q\rangle - \langle y, Q_{xx}y + Q_{xs}s + q\rangle - \tau - \lambda\|(s,x) - (\widehat{s}_i, \widehat{x}_i)\|_2^2 \\ \text{s.t.} & Cs \succeq_{\mathcal{C}} d,\ Wx \succeq_{\mathcal{K}} Hs + h,\ Wy \succeq_{\mathcal{K}} Hs + h \end{array} \tag{27}$$

does not exceed r_i. The Lagrangian of the conic quadratic program (27) is defined as

$$\begin{aligned} &L(s, x, y; \phi_{i1}, \mu_{i1}, \gamma_i) \\ &\quad := \langle x, Q_{xx}x + Q_{xs}s + q\rangle - \langle y, Q_{xx}y + Q_{xs}s + q\rangle - \tau - \lambda\|(s,x) - (\widehat{s}_i, \widehat{x}_i)\|_2^2 \\ &\qquad + \langle Cs - d, \phi_{i1}\rangle + \langle Wx - Hs - h, \mu_{i1}\rangle + \langle Wy - Hs - h, \gamma_i\rangle, \end{aligned} \tag{28}$$

and therefore (27) can be expressed as a max-min problem of the form

$$\begin{aligned} &\sup_{s,x,y} \inf_{\phi_{i1}\in\mathcal{C}^*} \inf_{\mu_{i1},\gamma_i\in\mathcal{K}^*} L(s, x, y; \phi_{i1}, \mu_{i1}, \gamma_i) \\ &\quad \le \inf_{\phi_{i1}\in\mathcal{C}^*} \inf_{\mu_{i1},\gamma_i\in\mathcal{K}^*} \sup_{s,x,y} L(s, x, y; \phi_{i1}, \mu_{i1}, \gamma_i), \end{aligned}$$

where the inequality follows from weak duality. As Ξ contains a Slater point by virtue of Assumption 5.1, strong duality holds (meaning that the inequality collapses to an equality) if the Lagrangian is concave in (s, x, y); see also Proposition 6.4 below. We conclude that the i-th member of the first constraint group in (26) holds if there exist $\phi_{i1} \in \mathcal{C}^*$ and $\mu_{i1}, \gamma_i \in \mathcal{K}^*$ with $\sup_{s,x,y} L(s, x, y; \phi_{i1}, \mu_{i1}, \gamma_i) \le r_i$. As the Lagrangian constitutes a quadratic function, this statement is satisfied if and only if there are $\phi_{i1} \in \mathcal{C}^*$, $\mu_{i1}, \gamma_i \in \mathcal{K}^*$, $\chi_{i1} \in \mathbb{R}^m$, $\zeta_{i1}, \eta_{i1} \in \mathbb{R}^n$ and $\rho_{i1} \in \mathbb{R}$ with

$$\begin{aligned} &\chi_{i1} = \tfrac{1}{2}(-C^\intercal\phi_{i1} + H^\intercal(\mu_{i1} + \gamma_i) - 2\lambda\widehat{s}_i) \\ &\zeta_{i1} = \tfrac{1}{2}(-q - W^\intercal\mu_{i1} - 2\lambda\widehat{x}_i), \quad \eta_{i1} = \tfrac{1}{2}(q - W^\intercal\gamma_i) \\ &\rho_{i1} = \tau + r_i + \lambda\left(\langle\widehat{x}_i, \widehat{x}_i\rangle + \langle\widehat{s}_i, \widehat{s}_i\rangle\right) + \langle d, \phi_{i1}\rangle + \langle h, \mu_i + \gamma_i\rangle \\ &\begin{bmatrix} \lambda\mathbb{I} & -\frac{1}{2}Q_{xs}^\intercal & \frac{1}{2}Q_{xs}^\intercal & \chi_{i1} \\ -\frac{1}{2}Q_{xs} & \lambda\mathbb{I} - Q_{xx} & 0 & \zeta_{i1} \\ \frac{1}{2}Q_{xs} & 0 & Q_{xx} & \eta_{i1} \\ \chi_{i1}^\intercal & \zeta_{i1}^\intercal & \eta_{i1}^\intercal & \rho_{i1} \end{bmatrix} \succeq 0. \end{aligned}$$

Similarly, it can be shown that the i-th member of the second constraint group in (26) is satisfied if and only if there exist $\phi_{i2} \in \mathcal{C}^*$, $\mu_{i2} \in \mathcal{K}^*$, $\chi_{i2} \in \mathbb{R}^m$, $\zeta_{i2} \in \mathbb{R}^n$ and $\rho_{i2} \in \mathbb{R}$ such that

$$\begin{aligned} &\chi_{i2} = \tfrac{1}{2}(-C^\intercal\phi_{i2} + H^\intercal\mu_{i2} - 2\lambda P_s\widehat{s}_i) \\ &\zeta_{i2} = \tfrac{1}{2}(-W^\intercal\mu_{i2} - 2\lambda\widehat{x}_i) \\ &\rho_{i2} = r_i + \lambda\left(\langle\widehat{x}_i, \widehat{x}_i\rangle + \langle\widehat{s}_i, \widehat{s}_i\rangle\right) + \langle d, \phi_{i2}\rangle + \langle h, \mu_i\rangle \\ &\begin{bmatrix} \lambda\mathbb{I} & 0 & \chi_{i2} \\ 0 & \lambda\mathbb{I} & \zeta_{i2} \\ \chi_{i2}^\intercal & \zeta_{i2}^\intercal & \rho_{i2} \end{bmatrix} \succeq 0. \end{aligned} \tag{29}$$

Replacing the semi-infinite constraints in (26) with the respective semidefinite approximations shows that the worst-case expectation (25) is bounded above by the optimal value of the conic program

$$
\begin{array}{ll}
\inf\ \varepsilon^2\lambda + \frac{1}{N}\sum_{i=1}^{N} r_i & \\
\text{s.t. } \lambda \in \mathbb{R}_+,\ r_i, \rho_{i1}, \rho_{i2} \in \mathbb{R},\ \phi_{i1}, \phi_{i2} \in \mathcal{C}^*,\ \mu_{i1}, \mu_{i2}, \gamma_i \in \mathcal{K}^* & \forall i \le N \\
\quad \chi_{i1}, \chi_{i2} \in \mathbb{R}^m,\ \zeta_{i1}, \eta_{i1}, \zeta_{i2} \in \mathbb{R}^n & \forall i \le N \\
\quad \chi_{i1} = \frac{1}{2}(-C^\intercal\phi_{i1} + H^\intercal(\mu_{i1} + \gamma_i) - 2\lambda\widehat{s}_i) & \forall i \le N \\
\quad \zeta_{i1} = \frac{1}{2}(-q - W^\intercal\mu_{i1} - 2\lambda\widehat{x}_i), \quad \eta_{i1} = \frac{1}{2}(q - W^\intercal\gamma_i) & \forall i \le N \\
\quad \rho_{i1} = \tau + r_i + \lambda\left(\langle\widehat{x}_i, \widehat{x}_i\rangle + \langle\widehat{s}_i, \widehat{s}_i\rangle\right) + \langle d, \phi_{i1}\rangle + \langle h, \mu_{i1} + \gamma_i\rangle & \forall i \le N \\
\quad \chi_{i2} = \frac{1}{2}(-C^\intercal\phi_{i2} + H^\intercal\mu_{i2} - 2\lambda\widehat{s}_i) & \forall i \le N \\
\quad \zeta_{i2} = \frac{1}{2}(-W^\intercal\mu_{i2} - 2\lambda\widehat{x}_i) & \forall i \le N \\
\quad \rho_{i2} = r_i + \lambda\left(\langle\widehat{x}_i, \widehat{x}_i\rangle + \langle\widehat{s}_i, \widehat{s}_i\rangle\right) + \langle d, \phi_{i2}\rangle + \langle h, \mu_i\rangle & \forall i \le N \\
\quad \begin{bmatrix} \lambda\mathbb{I} & -\frac{1}{2}Q_{xs}^\intercal & \frac{1}{2}Q_{xs}^\intercal & \chi_{i1} \\ -\frac{1}{2}Q_{xs} & \lambda\mathbb{I} - Q_{xx} & 0 & \zeta_{i1} \\ \frac{1}{2}Q_{xs} & 0 & Q_{xx} & \eta_{i1} \\ \chi_{i1}^\intercal & \zeta_{i1}^\intercal & \eta_{i1}^\intercal & \rho_{i1} \end{bmatrix} \succeq 0, \quad \begin{bmatrix} \lambda\mathbb{I} & 0 & \chi_{i2} \\ 0 & \lambda\mathbb{I} & \zeta_{i2} \\ \chi_{i2}^\intercal & \zeta_{i2}^\intercal & \rho_{i2} \end{bmatrix} \succeq 0 & \forall i \le N. \\
\end{array}
\tag{30}
$$

The claim then follows by substituting (30) into a suitable worst-case CVaR formula akin to (19). □

If $(\widehat{s}_i, \widehat{x}_i) \in \Xi$ for all $i \le N$, then the conic program (24) simplifies. In this case, the maximization problems in the last constraint group of (26) all evaluate to zero, which implies that the constraints (29) reduce to $r_i \ge 0$, while the decision variables ϕ_{i2}, μ_{i2}, χ_{i2}, ζ_{i2} and ρ_{i2} can be omitted from (24) for all $i \le N$.

In spite of the hardness results outlined in Theorems 6.1 and 6.2, the following proposition shows that the tractable approximation of Theorem 6.3 is sometimes exact.

Proposition 6.4 (Ex post exactness guarantee) *If an optimal solution to the safe conic approximation* (24) *from Theorem* 6.3 *satisfies the strict inequality*

$$
\begin{bmatrix} \lambda\mathbb{I} & -\frac{1}{2}Q_{xs}^\intercal & \frac{1}{2}Q_{xs}^\intercal \\ -\frac{1}{2}Q_{xs} & \lambda\mathbb{I} - Q_{xx} & 0 \\ \frac{1}{2}Q_{xs} & 0 & Q_{xx} \end{bmatrix} \succ 0, \tag{31}
$$

then (24) *is equivalent to the original distributionally robust inverse optimization problem* (24).

Our computational experiments suggest that the ex post exactness condition (31) is often satisfied if the Wasserstein radius ε is not too large.

Proof of Proposition 6.4 From the proof of Theorem 6.3 we conclude that the optimal value of the distributionally robust inverse optimization problem (13) can be represented as $\inf_{\theta,\lambda,\tau} \varphi(\theta, \lambda, \tau)$, where

$$\varphi(\theta, \lambda, \tau) := \lambda\varepsilon^2 + \frac{1}{N}\sum_{i=1}^{N} \sup_{s\in\mathbb{S},\, x,y\in\mathbb{X}(s)} \max\Big\{\big\langle x,\, Q_{xx}x + Q_{xs}s + q\big\rangle\big\langle y,\, Q_{xx}y + Q_{xs}s + q\big\rangle - \tau, 0\Big\} - \lambda\|(s,x) - (\widehat{s}_i, \widehat{x}_i)\|_2^2$$

for $\theta \in \Theta$ and $\lambda \geq 0$, and $\varphi(\theta, \lambda, \tau) = \infty$ otherwise. By construction, φ is jointly convex in $\theta = (Q_{xx}, Q_{xs}, q)$, λ and τ. Moreover, it is clear that the optimal value of the finite convex program (24) can be represented as $\inf_{\theta,\lambda,\tau} \widehat{\varphi}(\theta, \lambda, \tau)$, where $\widehat{\varphi}(\theta, \lambda, \tau)$ denotes the infimum of (24) when the decision variables θ, λ and τ are fixed. The proof of Theorem 6.3 further implies that $\widehat{\varphi}$ coincides with φ whenever the Lagrangian (28) is concave in (s, x, y), which is the case if and only if

$$\begin{bmatrix} -\lambda\mathbb{I} & \frac{1}{2}Q_{xs}^\mathsf{T} & -\frac{1}{2}Q_{xs}^\mathsf{T} \\ \frac{1}{2}Q_{xs} & Q_{xx} - \lambda\mathbb{I} & 0 \\ -\frac{1}{2}Q_{xs} & 0 & -Q_{xx} \end{bmatrix} \preceq 0. \tag{32}$$

Note also that $\widehat{\varphi}(\theta, \lambda, \tau) = \infty$ whenever (32) is violated because (32) is implied by the constraints of (24). In summary, φ and $\widehat{\varphi}$ are both convex functions satisfying

$$\widehat{\varphi}(\theta, \lambda, \tau) = \begin{cases} \varphi(\theta, \lambda, \tau) & \text{if (32) holds,} \\ +\infty & \text{otherwise.} \end{cases}$$

Thus, any minimizer of $\widehat{\varphi}$ satisfying the strict inequality (31) resides in the interior of the region where the convex functions $\widehat{\varphi}$ and φ coincide and must therefore also minimize φ.

Finally, we generalize Theorem 6.3 to the bounded rationality loss (10). The proof largely parallels that of Theorem 6.3 and is therefore omitted for brevity.

Corollary 6.5 (Quadratic hypotheses and bounded rationality loss) *Assume that $\mathcal{F}$ represents the class of quadratic hypotheses with a search space of the form (22) and that Assumption 5.1 holds. If the observer uses the bounded rationality loss (10) and measures risk using the* CVaR *at level $\alpha \in (0, 1]$, then the inverse optimization problem (13) over the 2-Wasserstein ball is is conservatively approximated by a variant of the conic program (24) where $\tau \in \mathbb{R}_+$ and the defining equation of ρ_i is replaced with*

$$\rho_{i1} = \tau + r_i + \delta + \lambda\left(\langle\widehat{x}_i, \widehat{x}_i\rangle + \langle\widehat{s}_i, \widehat{s}_i\rangle\right) + \langle d, \phi_{i1}\rangle + \langle h, \mu_{i1} + \gamma_i\rangle \quad \forall i \leq N.$$

The distributionally robust inverse optimization problem (13) also admits a safe convex approximation when the search space consists of a class of convex hypotheses of the type considered in [27]. Due to space restrictions we relegate this discussion to Appendix 1.

7 Numerical experiments

We now compare the proposed distributionally robust approach to inverse optimization against the state-of-the-art techniques described in [6] and [10]. The first experiment aims to learn a linear hypothesis from imperfect training samples, where the imperfection is explained by measurement noise or the agent's bounded rationality. Similarly, the second experiment endeavors to learn a quadratic hypothesis from imperfect training samples, where the imperfection is explained by measurement noise or model uncertainty.

All experiments are run on an Intel XEON CPU with 3.40GHz clock speed and 16GB of RAM. All linear, quadratic and second-order cone programs are solved with CPLEX 12.6, and all semidefinite programs are solved with MOSEK 8. In order to ensure that our experiments are reproducible, we make the underlying source codes available at https://github.com/sorooshafiee/InverseOptimization.

7.1 Learning a linear hypothesis

The goal of this experiment is to learn a linear hypothesis from imperfect training samples where the measured responses represent feasible perturbations of the true optimal responses. The perturbations can be explained by measurement noise or by the agent's bounded rationality. We argue that different causes of the perturbations necessitate different inverse optimization models.

7.1.1 Consistent noisy measurements

Decision problem of the agent: We assume that the agent's true objective function is $F(s, x) = \langle \theta^\star, x \rangle$ for some $\theta^\star$ in the vicinity of a nominal model θ_0. The nominal model is sampled uniformly from $\Theta_0 := \left\{\theta \in \mathbb{R}^n : \|\theta_0\|_\infty \in [1, 5]\right\}$, while $\theta^\star$ is sampled uniformly form $\Theta := \left\{\theta \in \mathbb{R}^n : \|\theta - \theta_0\|_\infty \leq 1\right\}$. This construction implies that $\theta^\star \neq 0$ almost surely. We also assume that the agent's feasible set is given by $\mathbb{X}(s) = \{x \in \mathbb{R}^n : \|x\|_\infty \leq 1,\ Ax \geq s\}$, where the constraint matrix A is sampled uniformly from the hypercube $[-1, 1]^{m \times n}$. This feasible set can be brought to the standard form of Assumption 5.1 by setting

$$W = (\mathbb{I}, -\mathbb{I}, A^\intercal)^\intercal \in \mathbb{R}^{(2n+m)\times n},\ H = (0, 0, \mathbb{I})^\intercal \in \mathbb{R}^{(2n+m)\times m}, h = (-1, -1, 0)^\intercal \in \mathbb{R}^{2n+m} \text{ and } \mathcal{K} = \mathbb{R}^{2n+m}_+.$$

For a fixed signal s, we denote the optimal value of the agent's true decision problem by $z^\star(s)$.

Generation of training samples: A single signal is constructed as $s = Av_1$, where the auxiliary random variable v_1 follows the uniform distribution on $[-1, 1]^n$. Thus, if a_i, $i \leq m$, denote the rows of the constraint matrix A, then the support of s can be expressed as $\mathbb{S} = \{s \in \mathbb{R}^m : |s_i| \leq \|a_i\|_1\ \forall i \leq m\}$. This support can be brought to the standard form of Assumption 5.1 by setting

$$C = (\mathbb{I}, -\mathbb{I})^\intercal \in \mathbb{R}^{2m \times m}, d = (\|a_1\|_1, \ldots, \|a_m\|_1, \|a_1\|_1, \ldots, \|a_m\|_1)^\intercal \in \mathbb{R}^{2m} \text{ and } \mathcal{C} = \mathbb{R}^{2m}_+.$$

As $v_1 \in \mathbb{X}(s)$ by construction, problem (1) is feasible for every $s \in \mathbb{S}$. We assume that the agent's response to s is noisy in the sense that it constitutes a random δ-suboptimal solution to (1) for $\delta = 1$. Specifically, we assume that x is obtained as a solution of an auxiliary optimization problem

$$\min_{x \in \mathbb{X}(s)} \left\{ \langle \theta_{\text{rand}}, x \rangle : \langle \theta^\star, x \rangle \leq z^\star(s) + \delta \right\},$$

which minimizes a linear cost over the set of all δ-suboptimal solutions to (1). The gradient θ_{rand} of the cost is sampled uniformly from $[-1, 1]^n$. By construction, we have $x \in \mathbb{X}(s)$, which implies that $(s, x) \in \Xi$. Thus, the measurement noise is consistent with the known support of the exact signal-response pairs. The imperfect consistent training samples $(\widehat{s}_i, \widehat{x}_i)$, $i \leq N$, are now generated independently using the above procedure.

Decision problem of the observer: The observer aims to identify a linear hypothesis $F_\theta(s, x) := \langle \theta, x \rangle$, $\theta \in \Theta$, that best predicts the agent's responses to new signals, where the search space Θ is set to the ∞-norm ball around the nominal model θ_0 described above. We assume that the observer minimizes the suboptimality loss (4b) and uses the expected value to measure risk. Moreover, the observer solves the distributionally robust inverse optimization problem (13) over a 1-Wasserstein ball around the empirical distribution on the training samples, where the ∞-norm is used as the transportation cost on Ξ. Note that this problem can be reformulated as the tractable linear program (18) by virtue of Theorem 5.2.

Out-of-sample risk: To assess the quality of an estimator $\widehat{\theta}_N$ obtained from (18), we evaluate its out-of-sample risk $\mathbb{E}^{\mathbb{P}_{\text{out}}}(\ell_{\widehat{\theta}_N})$ both with respect to the predictability loss (4a) and the suboptimality loss (4b), where $\mathbb{P}_{\text{out}}$ represents the distribution of a single test sample (s, x) independent of the training samples, and where x is an exact (non-noisy) response to s in (1). In practice, the out-of-sample risk cannot be evaluated analytically but only numerically by using 1000 independent test samples from $\mathbb{P}_{\text{out}}$. By relying on non-noisy test samples, we assess how well the estimator $\widehat{\theta}_N$ can predict the agent's *exact* optimal responses.

Results: All numerical results are averaged across 100 independent problems instances $\{\theta_0, \theta^\star, A\}$. The first experiment involves $m = 50$ signal variables, $n = 50$ decision variables and $N = 10$ training samples. Figure 1a shows how the out-of-sample risk of the optimal estimator $\widehat{\theta}_N(\varepsilon)$ obtained from (18) changes with the radius ε of the underlying Wasserstein ball. This experiment suggests that both the predictability and suboptimality risks can be significantly reduced by using a distributionally robust inverse optimization model with a judiciously calibrated ambiguity set. Unfortunately, Fig. 1a, from which the optimal Wasserstein radii could be read off easily, is not available in the training phase as its construction requires large amounts of test samples. Instead, the Wasserstein radii offering the lowest out-of-sample risk must also be esti-

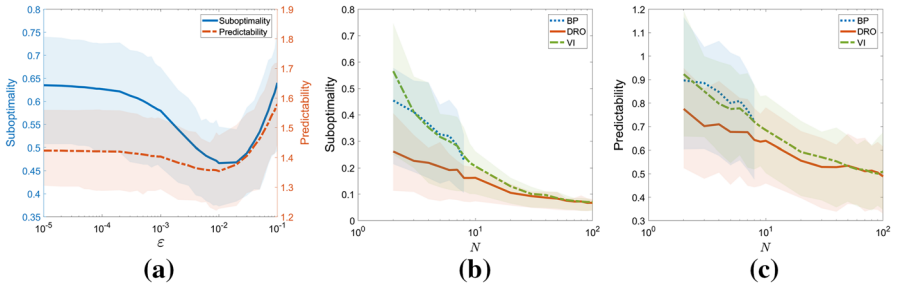

Fig. 1 Out-of-sample suboptimality and predictability risks in the presence of imperfect consistent training samples and perfect test samples. Lines represent averages across 100 simulation runs, while shaded areas visualize the tubes between the 25 and 75% quantiles. **a** Impact of the Wasserstein radius on the suboptimality and predictibility risk, **b** suboptimality learning curve, **c** predictibility learning curve

mated from the training data. To this end, we use the following *k-fold cross validation* algorithm:

Partition $\widehat{\xi}_1, \ldots, \widehat{\xi}_N$ into k folds, and repeat the following procedure for each fold $i = 1, \ldots, k$. Use the i-th fold as a validation dataset and merge the remaining $k-1$ folds to a training dataset of size N_i. Using only the i-th training dataset, solve (18) for a large but finite number of candidate radii $\varepsilon \in \mathcal{E}$ to obtain an estimator $\widehat{\theta}_{N_i}(\varepsilon)$. Use the i-th validation dataset to estimate the out-of-sample risk of $\widehat{\theta}_{N_i}(\varepsilon)$ for each $\varepsilon \in \mathcal{E}$. Set $\widehat{\varepsilon}_N^i$ to any $\varepsilon \in \mathcal{E}$ that minimizes this quantity. After all folds have been processed, set $\widehat{\varepsilon}_N$ to the average of the $\widehat{\varepsilon}_N^i$ for $i = 1, \ldots, k$, and re-solve (18) with $\varepsilon = \widehat{\varepsilon}_N$ using all N samples. Report the optimal solution $\widehat{\theta}_N$ and the optimal value $\widehat{J}_N$ of (18) as the recommended estimator and its corresponding certificate, respectively. Throughout all experiments we set $k = \min\{5, N\}$ and $\mathcal{E} = \{\varepsilon = b \cdot 10^c : b \in \{1, 5\},\ c \in \{-4, -3, -2, -1\}\}$.

For brevity, we henceforth refer to the above cross-validation scheme as the *distributionally robust optimization* (DRO) approach. In the following, we compare the resulting DRO estimator against two state-of-the-art estimators from the literature. The first estimator is obtained from the *variational inequality* (VI) approach proposed in [10], which minimizes the first-order loss (4c) averaged across all training samples. By [10, Theorem 3], the resulting inverse optimization problem is equivalent to the tractable linear program

$$\begin{array}{lll} \text{minimize} & \dfrac{1}{N}\displaystyle\sum_{i=1}^{N} |r_i| & \\ \text{subject to} & \theta \in \Theta,\ \gamma_i \in \mathbb{R}_+,\ r_i \in \mathbb{R} & \forall i \leq N \\ & \left\langle W\widehat{x}_i - H\widehat{s}_i - h, \gamma_i \right\rangle \leq r_i & \forall i \leq N \\ & W^\top \gamma_i = \theta & \forall i \leq N. \end{array}$$

Note that as all training samples are consistent, that is, $(\widehat{s}_i, \widehat{x}_i) \in \Xi$ for $i = 1, \ldots, n$, one can show that all residuals r_i are automatically non-negative, and the above linear program can be viewed as a special instance of (13) that minimizes the first-order loss, where the risk measure is set to the expected value and the Wasserstein radius is set

to zero. If there were inconsistent training samples that fall outside of Ξ, on the other hand, the absolute values of the residuals would be needed to prevent unboundedness.

As the first-order loss coincides with the suboptimality loss when focusing on linear hypotheses (see Sect. 5), the VI approach can be viewed as a special case of the DRO approach when all training samples are consistent *and* the Wasserstein radius ε is set to zero. Due to the extra freedom of choosing a nonzero ε, we therefore expect the DRO approach to have an advantage over the VI approach in this setting.

The second estimator is obtained from the *bilevel programming* (BP) approach proposed in [6], which minimizes the predictability loss (4a) averaged across all training samples. As shown in [6, Section 2.2], the resulting inverse optimization problem is equivalent to the optimistic bilevel program

$$\begin{array}{ll} \text{minimize} & \dfrac{1}{N}\displaystyle\sum_{i=1}^{N} \|\widehat{x}_i - y_i\|_2^2 \\ \text{subject to} & \theta \in \Theta,\ \ y_i \in \arg\min\limits_{z} \left\{ \langle z, \theta \rangle \ :\ Wz \geq H\widehat{s}_i + h \right\}\ \forall i \leq N. \end{array}$$

Note that this bilevel program can be viewed as a special instance of (13) that minimizes the predictability loss, where the risk measure is set to the expected value and the Wasserstein radius is set to zero.

The above bilevel program can be reformulated as a mixed integer quadratic program by replacing the 'arg min'-constraint with the Karush–Kuhn–Tucker optimality conditions of the agent's linear program. Indeed, note that the resulting complementary slackness conditions can be linearized by introducing binary variables that identify the binding constraints. However, this approach requires big-M constants to bound the optimal dual variables. As valid big-M constants are difficult to obtain in general and as overly conservative big-M constants lead to numerical instability, we use here the YALMIP interface for bilevel programming, which calls a dedicated branch and bound algorithm that branches directly on the complementarity slackness conditions [28]. Throughout our experiments we limit all branch and bound calculations to 2000 iterations.

Figure 1b, c visualize the suboptimality and predictability *learning curves*, respectively, which capture how the out-of-sample risk of the different estimators changes with the number of training samples. As optimistic bilevel programs are NP-hard even when all objective and constraint functions are linear [8], this experiment focuses on problem instances with $n = m = 10$. Even for these moderate problem dimensions, however, the inverse optimization problems associated with the BP approach fails to find a feasible solution for more than 10 training samples. In contrast, all other approaches lead to tractable linear programs. Figure 1b shows that the DRO approach dominates the VI and BP approaches uniformly across all samples sizes in terms of out-of-sample suboptimality risk. This is reassuring as the DRO approach actually minimizes suboptimality risk. However, Fig. 1c suggests that the DRO approach wins even in terms of out-of-sample predictability risk. This is perhaps surprising because, unlike the BP approach, the DRO approach only minimizes an approximation of the predictability loss. We conclude that injecting distributional robustness may be more beneficial for out-of-sample performance than using the correct loss function.

Table 1 Out-of-sample suboptimality risk in the presence of noisy consistent measurements

		m				
n	Methods	10	20	30	40	50
10	VI	1.4×10^{-1}	1.8×10^{-1}	1.4×10^{-1}	1.2×10^{-1}	1.3×10^{-1}
	DRO	1.2×10^{-1}	1.3×10^{-1}	1.1×10^{-1}	8.9×10^{-2}	9.5×10^{-2}
20	VI	3.0×10^{-1}	3.5×10^{-1}	3.2×10^{-1}	2.5×10^{-1}	2.4×10^{-1}
	DRO	2.4×10^{-1}	2.9×10^{-1}	2.4×10^{-1}	1.9×10^{-1}	1.9×10^{-1}
30	VI	3.8×10^{-1}	4.4×10^{-1}	4.6×10^{-1}	4.2×10^{-1}	3.6×10^{-1}
	DRO	3.2×10^{-1}	3.7×10^{-1}	3.7×10^{-1}	3.3×10^{-1}	3.1×10^{-1}
40	VI	4.4×10^{-1}	5.5×10^{-1}	6.0×10^{-1}	5.4×10^{-1}	5.1×10^{-1}
	DRO	3.7×10^{-1}	4.5×10^{-1}	4.6×10^{-1}	4.4×10^{-1}	4.5×10^{-1}
50	VI	5.0×10^{-1}	6.4×10^{-1}	6.5×10^{-1}	7.1×10^{-1}	6.2×10^{-1}
	DRO	4.1×10^{-1}	5.1×10^{-1}	5.4×10^{-1}	5.6×10^{-1}	5.4×10^{-1}

Table 2 Out-of-sample predictability risk in the presence of consistent noisy measurements

		m				
n	Methods	10	20	30	40	50
10	VI	6.1×10^{-1}	4.6×10^{-1}	3.4×10^{-1}	2.6×10^{-1}	2.4×10^{-1}
	DRO	5.6×10^{-1}	4.2×10^{-1}	3.2×10^{-1}	2.4×10^{-1}	2.2×10^{-1}
20	VI	1.1×10^{0}	9.1×10^{-1}	7.6×10^{-1}	6.1×10^{-1}	5.1×10^{-1}
	DRO	1.0×10^{0}	9.0×10^{-1}	7.2×10^{-1}	5.9×10^{-1}	4.8×10^{-1}
30	VI	1.4×10^{0}	1.2×10^{0}	1.1×10^{0}	9.7×10^{-1}	8.5×10^{-1}
	DRO	1.4×10^{0}	1.2×10^{0}	1.1×10^{0}	9.5×10^{-1}	8.4×10^{-1}
40	VI	1.6×10^{0}	1.5×10^{0}	1.4×10^{0}	1.3×10^{0}	1.2×10^{0}
	DRO	1.6×10^{0}	1.5×10^{0}	1.4×10^{0}	1.3×10^{0}	1.2×10^{0}
50	VI	1.8×10^{0}	1.7×10^{0}	1.6×10^{0}	1.6×10^{0}	1.4×10^{0}
	DRO	1.8×10^{0}	1.7×10^{0}	1.6×10^{0}	1.6×10^{0}	1.4×10^{0}

Tables 1 and 2 report the out-of-sample suboptimality and predictability risks, respectively, for the DRO and VI estimators based on $N = 10$ training samples and for different signal and response dimensions. The BP approach is excluded from this comparison due to its intractability. We observe that the DRO estimator always attains the lowest suboptimality risk and often the lowest predictability risk. Whenever the VI estimator wins, the predictability risks of the VI and DRO estimators are in fact almost identical.

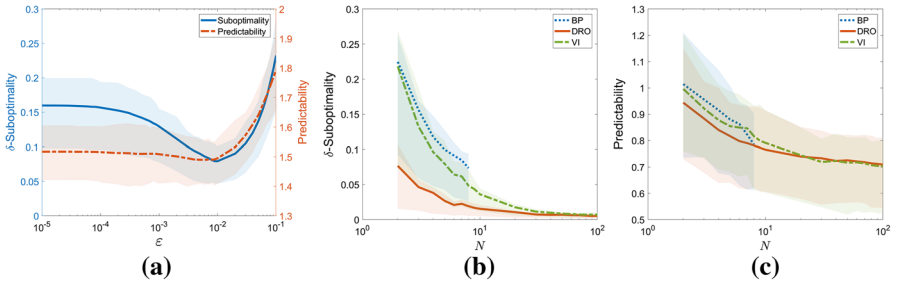

Fig. 2 Out-of-sample suboptimality and predictability risks in the presence of imperfect consistent training and test samples. Lines represent averages across 100 simulation runs, while shaded areas visualize the tubes between the 25 and 75% quantiles. **a** Impact of the Wasserstein radius on the suboptimality and predictibility risk, **b** suboptimality learning curve, **c** predictibility learning curve

7.1.2 Bounded rationality

Consider the exact same setting as in Sect. 7.1.1 but assume now that the agent suffers from bounded rationality. Specifically, assume that the agent selects random δ-suboptimal decisions that are perfectly measured by the observer. This means that the training samples are generated as in Sect. 7.1.1, but the imperfection of the training responses is now explained by the agent's bounded rationality instead of the observer's noisy measurements. At this point one may wonder why a new interpretation of the imperfections should impact the observer's inverse optimization problem. In the following we will assume, however, that the observer is aware of the agent's bounded rationality, knows the value of δ and aims to predict the agent's actual suboptimal decisions. Therefore, the DRO approach to inverse optimization is subject to two changes:

- The observer minimizes the bounded rationality loss (10) instead of the suboptimality loss (4b).
- The test samples are generated in the same way as the training samples. Thus, the test responses no longer constitute perfect minimizers of (1) but represent random δ-suboptimal solutions.

Recall that the bounded rationality loss favors hypotheses that correctly predict δ-suboptimal responses. We also highlight that the observer's inverse optimization problem with bounded rationality loss can be reformulated as a tractable linear program by virtue of Corollary 5.4. Moreover, by using imperfect test samples, the out-of-sample risk is now measured under the distribution of an *imperfect* signal-response pair, which is the correct performance criterion given that the observer aims to predict imperfect responses.

In analogy to Sect. 7.1.1, the impact of the Wasserstein radius on the out-of-sample suboptimality and predictability risk is shown in Fig. 2a. The suboptimality and predictability learning curves of different estimators are visualized in Fig. 2b, c, respectively. Here, the DRO estimator is defined in terms of the bounded rationality loss, while the VI and BP estimators are computed as in Sect. 7.1.1 and are thus not corrected for the agent's bounded rationality. The impact of the signal and

Table 3 Out-of-sample suboptimality risk in the presence of bounded rationality

n	Methods	m = 10	20	30	40	50
10	VI	4.3×10^{-2}	2.5×10^{-2}	1.1×10^{-2}	5.7×10^{-3}	4.5×10^{-3}
	DRO	2.5×10^{-2}	1.1×10^{-2}	6.4×10^{-3}	2.9×10^{-3}	1.8×10^{-3}
20	VI	7.7×10^{-2}	7.8×10^{-2}	6.3×10^{-2}	3.7×10^{-2}	1.8×10^{-2}
	DRO	5.3×10^{-2}	4.4×10^{-2}	3.4×10^{-2}	2.1×10^{-2}	9.2×10^{-3}
30	VI	1.2×10^{-1}	1.3×10^{-1}	1.2×10^{-1}	8.9×10^{-2}	6.4×10^{-2}
	DRO	7.9×10^{-2}	7.8×10^{-2}	6.9×10^{-2}	5.2×10^{-2}	3.9×10^{-2}
40	VI	1.4×10^{-1}	1.9×10^{-1}	1.7×10^{-1}	1.4×10^{-1}	1.2×10^{-1}
	DRO	9.7×10^{-2}	1.2×10^{-1}	1.1×10^{-1}	9.5×10^{-2}	8.2×10^{-2}
50	VI	1.9×10^{-1}	2.0×10^{-1}	2.1×10^{-1}	2.0×10^{-1}	1.7×10^{-1}
	DRO	1.2×10^{-1}	1.3×10^{-1}	1.4×10^{-1}	1.4×10^{-1}	1.2×10^{-1}

Table 4 Out-of-sample predictability risk in the presence of bounded rationality

n	Methods	m = 10	20	30	40	50
10	VI	8.2×10^{-1}	5.8×10^{-1}	4.5×10^{-1}	3.5×10^{-1}	3.2×10^{-1}
	DRO	7.9×10^{-1}	5.9×10^{-1}	4.6×10^{-1}	3.7×10^{-1}	3.4×10^{-1}
20	VI	1.2×10^{0}	1.1×10^{0}	8.9×10^{-1}	7.2×10^{-1}	5.8×10^{-1}
	DRO	1.2×10^{0}	1.1×10^{0}	9.2×10^{-1}	7.4×10^{-1}	6.0×10^{-1}
30	VI	1.5×10^{0}	1.4×10^{0}	1.2×10^{0}	1.1×10^{0}	9.4×10^{-1}
	DRO	1.5×10^{0}	1.4×10^{0}	1.3×10^{0}	1.1×10^{0}	9.6×10^{-1}
40	VI	1.7×10^{0}	1.6×10^{0}	1.5×10^{0}	1.4×10^{0}	1.3×10^{0}
	DRO	1.8×10^{0}	1.7×10^{0}	1.6×10^{0}	1.4×10^{0}	1.3×10^{0}
50	VI	1.9×10^{0}	1.8×10^{0}	1.7×10^{0}	1.6×10^{0}	1.5×10^{0}
	DRO	1.9×10^{0}	1.9×10^{0}	1.8×10^{0}	1.7×10^{0}	1.6×10^{0}

response dimensions on the out-of-sample suboptimality and predictability risk are reported in Tables 3 and 4, respectively. Unless otherwise stated, all experiments are parameterized exactly as in Sect. 7.1.1. Here, the DRO estimator systematically attains the lowest suboptimality risk, while the VI estimator almost always wins in terms of predictability risk, even though only by a small margin.

7.2 Learning a quadratic hypothesis

A fundamental problem in marketing is to understand the purchasing decisions of consumers, which is an essential prerequisite for estimating demand functions. In this section we study the inverse optimization problem of a marketing manager (the observer) aiming to learn the quadratic utility function that best explains the purchasing decisions of a consumer (the agent). This problem setup is inspired by [27].

7.2.1 Consistent noisy measurements

Decision problem of the agent: Assume that there are n products with prices $s \in \mathbb{R}^n_+$. The agent aims to select a basket of goods $x \in \mathbb{R}^n_+$ that minimizes the true objective function $F(s,x) = \langle s, x\rangle - U(x)$, where $\langle s, x\rangle$ represents the purchasing costs, while the concave quadratic function $U(x) := -\langle x, Q^\star_{xx} x\rangle - \langle q^\star, x\rangle$ captures the utility of consumption. The positive definite matrix $Q^\star_{xx}$ is constructed as follows. Sample a square matrix A uniformly form $[-1,1]^{n\times n}$, denote by R the orthogonal matrix consisting of the orthonormal eigenvectors of $(A + A^\intercal)/2$ and set $Q^\star_{xx} := R^\intercal D R$, where D is a diagonal matrix whose main diagonal is sampled uniformly form $[0.2, 1]^n$. Moreover, the gradient $q^\star$ is sampled uniformly from $[-2, 0]^n$. Finally, define the agent's feasible set as $\mathbb{X}(s) = [0, 5]^n$, which can be brought to the standard form of Assumption 5.1 by setting

$$W = (\mathbb{I}, -\mathbb{I})^\intercal \in \mathbb{R}^{2n\times n},\ H = (0,0)^\intercal \in \mathbb{R}^{2n\times m},\ h = 5\cdot(0,-1)^\intercal \in \mathbb{R}^{2n} \text{ and } \mathcal{K} = \mathbb{R}^{2n}_+.$$

As usual, for a fixed signal s, we denote the optimal value of the agent's true decision problem by $z^\star(s)$.

Generation of training samples: Signals follow the uniform distribution on the support set $\mathbb{S} = [0, 1]^n$, which can be brought to the standard form of Assumption 5.1 by setting $C = (\mathbb{I}, -\mathbb{I})^\intercal \in \mathbb{R}^{2n\times n}$, $d = (0, -1)^\intercal \in \mathbb{R}^{2n}$ and $\mathcal{C} = \mathbb{R}^{2n}_+$. As in Sect. 7.1.1, we assume that the agent's response to s is noisy and constitutes a random δ-suboptimal solution to (1) for $\delta = 0.2$. Thus, x is obtained as a solution of the auxiliary problem

$$\min_{x\in\mathbb{X}(s)} \left\{ \langle x, Q_{\text{rand}} x\rangle : \langle x, Q^\star_{xx} x\rangle + \langle s, x\rangle + \langle q^\star, x\rangle \le z^\star(s) + \delta \right\},$$

which minimizes a convex quadratic cost over the set of all δ-suboptimal solutions to (1). Specifically, Q_{rand} is a diagonal matrix whose main diagonal is sampled uniformly form $[0, 1]^n$. As $(s, x) \in \Xi$ by construction, the measurement noise is consistent with the know support of the exact signal-response pairs. The imperfect consistent training samples $(\widehat{s}_i, \widehat{x}_i)$, $i \le N$, are now generated independently using the above procedure.

Decision problem of the observer: The observer aims to identify the best quadratic hypothesis of the form $F_\theta(s,x) := \langle x, Q_{xx} x\rangle + \langle x, s\rangle + \langle q, x\rangle$, where the parameter

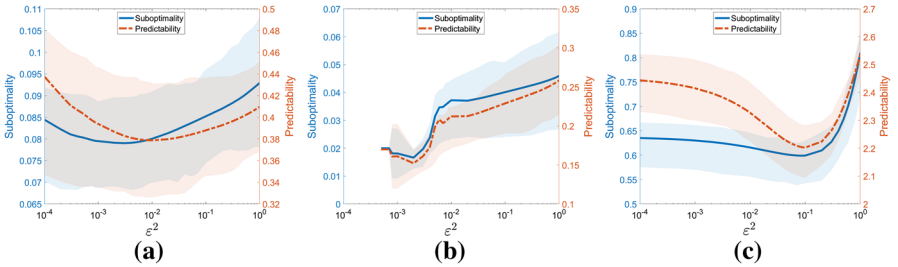

Fig. 3 Impact of the Wasserstein radius on the out-of-sample suboptimality and predictability risk. Lines represent averages across 100 simulation runs, while shaded areas visualize the tubes between the 25 and 75% quantiles. **a** Imperfect consistent training samples and perfect test samples, **b** imperfect inconsistent training samples and perfect test samples, **c** model uncertainty

$\theta = (Q_{xx}, q)$ ranges over the search space

$$\Theta = \Big\{\theta = (Q_{xx}, q) \in \mathbb{R}^{n\times n} \times \mathbb{R}^n \;:\; Q_{xx} \succeq 0\Big\}.$$

Note that no hypothesis $F_\theta(s, x)$, $\theta \in \Theta$, can vanish identically due to the term $\langle x, s\rangle$. Note also that the agent's true objective function corresponds to $\theta^\star = (Q^\star_{xx}, q^\star) \in \Theta$. We assume that the observer minimizes the suboptimality loss (4b), uses the expected value to measure risk and solves the distributionally robust inverse optimization problem (13) over a 2-Wasserstein ball around the empirical distribution on the training samples, where the 2-norm is used as the transportation cost on Ξ. By Theorem 6.3, the emerging inverse optimization problem is conservatively approximated by the tractable semidefinite program (24).

Out-of-sample risk: The quality of an estimator $\widehat{\theta}_N$ obtained from (24) is measured by its out-of-sample risk $\mathbb{E}^{\mathbb{P}_{\text{out}}}(\ell_{\widehat{\theta}_N})$ both with respect to the predictability loss (4a) and the suboptimality loss (4b), where $\mathbb{P}_{\text{out}}$ represents the distribution of a single test sample (s, x) independent of the training samples, and where x is an exact (non-noisy) response to s in (1). More precisely, the out-of-sample risk is evaluated approximately using 1,000 independent test samples from $\mathbb{P}_{\text{out}}$.

Results: All numerical results are averaged across 100 independent problems instances $\{Q^\star_{xx}, q^\star\}$. Under the assumption that the signal and response dimensions are set to $n = 10$ and there are $N = 20$ training samples, Fig. 3a shows how the out-of-sample risk of the optimal estimator $\widehat{\theta}_N(\varepsilon)$ obtained from (13) changes with the Wasserstein radius ε. As in Sect. 7.1.1, this experiment suggests that both the predictability and suboptimality risks can be reduced by using a distributionally robust inverse optimization model.

7.2.2 Inconsistent noisy measurements

Assume now that the agent's optimal solutions to (1) are corrupted by additive measurement noise that follows a uniform distribution on $[-0.1, 0.1]^n$. Otherwise, we consider the exact same experimental setup as in Sect. 7.2.1. Under this premise,

it is likely that some training responses are infeasible in (1), that is, $\widehat{x}_i \notin \mathbb{X}(\widehat{s}_i)$ and—*a fortiori*—$(\widehat{s}_i, \widehat{x}_i) \notin \Xi$ for some $i \leq N$. These problematic training samples are inconsistent with the known support of the perfect signal-response pairs, and the corresponding empirical distribution $\widehat{\mathbb{P}}_N$ fails to be supported on Ξ. Thus, for all sufficiently small values of ε there exists no distribution $\mathbb{Q}$ on Ξ with $\mathrm{W}_2\big(\mathbb{Q}, \widehat{\mathbb{P}}_N\big) \leq \varepsilon$, implying that the Wasserstein ball $\mathbb{B}_\varepsilon^2(\widehat{\mathbb{P}}_N)$ is empty, in which case the distributionally robust inverse optimization problem (24) becomes meaningless. Figure 3b visualizes the out-of-sample suboptimality and predictability risk of the optimal estimator $\widehat{\theta}(\varepsilon)$ as a function of ε. Note that for any fixed ε the figure reports the out-of-sample risk averaged only across those problem instances $\{Q_{xx}^\star, q^\star\}$ for which $\mathbb{B}_\varepsilon^2(\widehat{\mathbb{P}}_N) \neq \emptyset$. Inspecting the results at the instance level, we observe that the out-of-sample risk is typically minimized by the smallest value of $\varepsilon \geq 0$ for which $\mathbb{B}_\varepsilon^2(\widehat{\mathbb{P}}_N) \neq \emptyset$ (this is not evident from the aggregate results shown in Fig. 3b). We conclude that a distributionally robust approach may be necessary for consistency.

7.2.3 Model uncertainty

Assume next that the agent's objective function is not contained in the set of hypotheses $F_\theta(s, x)$, $\theta \in \Theta$, but that the signals and the agent's responses are unaffected by noise. Specifically, in analogy to [27], we assume that the true utility function is given by $U(x) := \big\langle 1, \sqrt{Ax - b} \big\rangle$, where A is a diagonal matrix whose main diagonal is sampled uniformly from $[0.5, 1]^n$, while b is sampled uniformly from $[0, 0.25]^n$. The square root is applied componentwise and evaluates to $-\infty$ for negative arguments. Otherwise, we consider the exact same experimental setup as in Sect. 7.2.1. Figure 3c shows the out-of-sample suboptimality and predictability risk of the optimal estimator $\widehat{\theta}(\varepsilon)$ as a function of ε, indicating that the best results are obtained for strictly positive Wasserstein radii, which enable the observer to combat over-fitting to the training samples.

7.2.4 Comparison of different data-driven inverse optimization schemes

In practice, the Wasserstein radii offering the lowest out-of-sample risk must be estimated from the training samples only. To this end, the DRO approach calibrates ε via k-fold cross validation as in Sect. 7.1. Another estimator is obtained by solving the *empirical risk minimization* (ERM) problem (8) using the empirical mean and the suboptimality loss (4b). This approach effectively mimics the DRO approach but sets $\varepsilon = 0$, which can be viewed as a trivial data-driven strategy to calibrate the Wasserstein radius.[4] The VI approach also disregards ambiguity and minimizes the empirical

[4] We did not consider the ERM approach in Sect. 5 because it coincides with the VI approach for linear hypotheses.

first-order loss by solving the semidefinite program

$$\begin{array}{lll}
\text{minimize} & \dfrac{1}{N}\displaystyle\sum_{i=1}^{N} |r_i| & \\
\text{subject to} & Q_{xx} \in \mathbb{R}^{n\times n},\ q \in \mathbb{R}^n,\ \gamma_i \in \mathbb{R}_+,\ r_i \in \mathbb{R} & \forall i \leq N \\
& \left\langle W\widehat{x}_i - H\widehat{s}_i - h, \gamma_i \right\rangle \leq r_i & \forall i \leq N \\
& W^\top \gamma_i = 2Q_{xx}\widehat{x}_i + \widehat{s}_i + q & \forall i \leq N \\
& Q_{xx} \succeq 0, &
\end{array}$$

see [10, Theorem 3]. As in Sect. 7.1.1, the absolute values of the residuals r_i in the objective can be dropped whenever the training samples are consistent with Ξ. We exclude the BP approach from this experiment because it leads to severely intractable mixed integer semidefinite programming problems.

All three approaches described above search over the *parametric* space of quadratic hypotheses. If the observer suspects that the true utility function fails to be quadratic, however, she may prefer a *non-parametric* approach that models the gradient of the utility function as a vector field $f \in \mathcal{H}^n$, where $\mathcal{H}$ represents a reproducing kernel Hilbert space of real-valued functions on $\mathbb{R}^n_+$, which is induced by a symmetric and positive definite kernel function $k : \mathbb{R}^n_+ \times \mathbb{R}^n_+ \to \mathbb{R}$. As $\mathcal{H}^n$ is typically infinite-dimensional and contains multiple candidate gradients $f \in \mathcal{H}^n$ that explain the training data, it has been suggested in [10, Section 5] to minimize the Hilbert norm $\sum_{i=1}^n \|f_i\|^2_{\mathcal{H}}$ of f subject to the constraint that the residuals of the first-order optimality condition at the training samples satisfy $\frac{1}{N}\sum_{i=1}^N |r_i| \leq \kappa$ for some prescribed threshold $\kappa \geq 0$. This amounts to finding the smoothest (with respect to the kernel function k) candidate gradient $f \in \mathcal{H}^n$ that explains the training data to within accuracy κ. By leveraging a generalized representer theorem, the resulting infinite-dimensional optimization problem can be reformulated as the tractable quadratic program

$$\begin{array}{lll}
\text{minimize} & \displaystyle\sum_{i=1}^{n} \left\langle e_i, \alpha K \alpha^\top e_i \right\rangle & \\
\text{subject to} & \alpha \in \mathbb{R}^{n\times N},\ \gamma_i \in \mathbb{R}_+,\ r_i \in \mathbb{R} & \forall i \leq N \\
& \left\langle W\widehat{x}_i - H\widehat{s}_i - h, \gamma_i \right\rangle \leq r_i & \forall i \leq N \\
& W^\top \gamma_i = \widehat{s}_i - \alpha K e_i & \forall i \leq N \\
& \dfrac{1}{N}\displaystyle\sum_{i=1}^{N} |r_i| \leq \kappa, &
\end{array} \tag{33a}$$

where $K \in \mathbb{R}^{N\times N}$ denotes the kernel matrix with entries $K_{ij} = k(\widehat{x}_i, \widehat{x}_j)$, e_i stands for the i-th standard basis vector in a space of appropriate dimension, and $\alpha \in \mathbb{R}^{n\times N}$ denotes a decision variable that encodes the partial derivatives of the utility function via $f_i(x) = \sum_{j=1}^N \alpha_{ij} k(\widehat{x}_j, x)$; see [10, Theorem 5].

In the inverse optimization context studied here, however, the above non-parametric VI approach suffers from two shortcomings that are not addressed in [10].

(i) The inverse optimization problem (33a) aims to learn the gradient field f of the unknown utility function U. As the Hessian matrix of U must be symmetric, the gradient field f must satisfy

$$\frac{\partial f_i(x)}{\partial x_j} = \frac{\partial f_j(x)}{\partial x_i} \qquad \forall x \in \mathbb{R}^n_+,\ \forall i, j = 1, \ldots, n.$$

By Stokes' theorem, this condition is necessary and sufficient to ensure that the utility function U can be reconstructed uniquely (modulo an additive constant) from f. Specifically, this condition guarantees that the utility function is defined unambiguously through the line integral $U(x) = U(0) + \int_C \langle f(x'), \mathrm{d}x' \rangle$, where C represents an arbitrary piecewise smooth curve in $\mathbb{R}^n_+$ connecting 0 and x.

(ii) While the inverse optimization problem (33a) represents a tractable quadratic program, the resulting gradient field f may induce a non-concave utility function U, which means that the agent's objective function $F(s, x) = \langle s, x \rangle - U(x)$ may have multiple local minima. Thus, even though learning f is easy, predicting the optimal response x to a given signal s may require the solution of a hard non-convex optimization problem, which may severely complicate extensive out-of-sample tests. Similarly, evaluating the suboptimality loss $F(s, x)$ at a fixed signal-response pair is intractable.

Here we address the first challenge by using the polynomial kernel function $k(x, x') = (c\langle x, x' \rangle + 1)^p$, which allows us to include the missing symmetry conditions in (33a) by appending the linear equality constraints

$$\sum_{k=1}^{N} \left(\alpha_{ik} [\widehat{x}_k]_j - \alpha_{jk} [\widehat{x}_k]_i \right) \prod_{t=1}^{q-1} [\widehat{x}_k]_{l_t} = 0 \quad \forall i, j = 1, \ldots, n, \forall 1 \leq l_1 \leq \cdots \leq l_{q-1} \leq n, \forall q = 1, \ldots, p. \tag{33b}$$

The second challenge does not have a simple remedy, and therefore we abandon the ideal goal to solve (1) to global optimality. Instead, for any given signal s, we use the gradient field f obtained from (33) directly to find a local solution of (1) via the classical subgradient descent algorithm [16, Chapter 3], where the step size is set to 10^{-3} and an initial feasible solution is selected uniformly at random from $\mathbb{X}(s)$. Note that the focus on local minima in prediction is consistent with the use of the first-order loss (4c) in inverse optimization because the first-oder loss cannot distinguish local and global optima and thus fails to penalize training samples $(\widehat{s}_i, \widehat{x}_i)$ where $\widehat{x}_i$ is a locally optimal but globally suboptimal response to $\widehat{s}_i$.

In our experiments we determine the parameter c of the polynomial kernel via 5-fold cross validation. Once the gradient field f has been inferred from (33), we construct the corresponding utility function by integrating f along the straight line C between 0 and x with parameterization $g(t) = tx$ for $t \in [0, 1]$, that is, we set

$$U(x) = U(0) + \int_C \langle f(x'), \mathrm{d}x' \rangle = U(0) + \int_0^1 \langle f(tx), x \rangle \mathrm{d}t = U(0)$$

Table 5 Data-driven inverse optimization: out-of-sample suboptimality and predictability risk of the VI, ERM and DRO approaches in different experimental settings

	Methods	Suboptimality	Predictability
Consistent noisy measurements	VI (parametric)	3.2×10^{-1}	1.3×10^{0}
	ERM	8.6×10^{-2}	4.7×10^{-1}
	DRO	7.9×10^{-2}	3.9×10^{-1}
Inconsistent noisy measurements	VI (parametric)	9.2×10^{-2}	6.3×10^{-1}
	ERM	Unbounded	Unbounded
	DRO	4.0×10^{-2}	2.4×10^{-1}
Model uncertainty	VI (parametric)	8.3×10^{-1}	2.8×10^{0}
	VI (non-parametric, $p = 2$)	1.9×10^{0}	4.8×10^{0}
	VI (non-parametric, $p = 3$)	4.0×10^{-1}	6.4×10^{0}
	ERM	6.2×10^{-1}	2.4×10^{0}
	DRO	6.0×10^{-1}	2.2×10^{0}

$$
\begin{aligned}
&+ \int_0^t \sum_{i=1}^{n} \sum_{j=1}^{N} x_i \alpha_{ij} k(\widehat{x}_j, tx)\, \mathrm{d}t \\
&= U(0) + \sum_{i=1}^{n} \sum_{j=1}^{N} x_i \alpha_{ij} \int_0^1 (c\langle tx, \widehat{x}_j \rangle + 1)^p \, \mathrm{d}t = U(0) \\
&+ \sum_{i=1}^{n} \sum_{j=1}^{N} x_i \alpha_{ij} \frac{\left(c\langle x, \widehat{x}_j \rangle + 1\right)^{p+1} - 1}{c(p+1)\langle x, \widehat{x}_j \rangle}.
\end{aligned}
$$

Table 5 reports the out-of-sample suboptimality and predictability risks, respectively, for the DRO, VI and ERM estimators. The non-parametric VI approach is exclusively used in the presence of model uncertainty, which is the only scenario in which it has a chance to outperform the more parsimonious parametric approaches. Our results show that the DRO estimator consistently attains the lowest suboptimality and predictability risk among all parametric approaches. We emphasize that the suboptimality and predicatability risk of the non-parametric VI approach are evaluated with respect to the local minimum identified by the subgradient descent algorithm and are therefore largely meaningless. In fact, we observed that the suboptimality risk becomes even negative for certain instances in which the global minimum of (1) could not be found. This artefact explains the low suboptimality risk of the non-parametric VI approach with $p = 3$. Note also that for $p \geq 4$ the number of symmetry conditions (33b) explodes, and thus (33) can no longer be solved. We also reconfirm our earlier observation that injecting robustness reduces out-of-sample risk.

Acknowledgements This work was supported by the Swiss National Science Foundation grant BSCGI0_157733.

Appendix A. Convex hypotheses

Consider the class $\mathcal{F}$ of hypotheses of the form $F_\theta(s,x) := \langle \theta, \Psi(x) \rangle$, where each component function of the feature map $\Psi : \mathbb{R}^n \to \mathbb{R}^d$ is convex, and where the weight vector θ ranges over a convex closed search space $\Theta \subseteq \mathbb{R}^d_+$. Thus, by construction, $F_\theta(s,x)$ is convex in x for every fixed $\theta \in \Theta$. In the remainder we will assume without much loss of generality that the transformation from the signal-response pair (s,x) to the signal-feature pair $(s, \Psi(x))$ is Lipschitz continuous with Lipschitz modulus 1, that is, we require

$$\|(s_1, \Psi(x_1)) - (s_2, \Psi(x_2))\| \leq \|(s_1, x_1) - (s_2, x_2)\| \quad \forall (s_1, x_1), (s_2, x_2) \in \mathbb{R}^m \times \mathbb{R}^n. \tag{34}$$

Before devising a safe conic approximation for the inverse optimization problem (13) with convex hypotheses, we recall that the conjugate of a function $f : \mathbb{R}^n \to \mathbb{R}$ is defined through $f^*(z) = \sup_{x \in \mathbb{R}^n} \langle z, x \rangle - f(x)$.

Theorem A.1 (Convex hypotheses and suboptimality loss) *Assume that $\mathcal{F}$ represents the class of convex hypotheses induced by the feature map Ψ and with a convex closed search space $\Theta \subseteq \mathbb{R}^d_+$ and that Assumption 5.1 holds. If the observer uses the suboptimality loss (4b) and measures risk using the* CVaR *at level $\alpha \in (0,1]$, then the following convex program provides a safe approximation for the distributionally robust inverse optimization problem (13) over the 1-Wasserstein ball:*

$$\begin{array}{lll}
\text{minimize} & \tau + \dfrac{1}{\alpha}\left(\varepsilon\lambda + \dfrac{1}{N}\displaystyle\sum_{i=1}^N r_i\right) & \\
\text{subject to} & \theta \in \Theta,\ \ \lambda \in \mathbb{R}_+,\ \ \tau, r_i \in \mathbb{R},\ \ \phi_{i1}, \phi_{i2} \in \mathcal{C}^*,\ \ \gamma_i \in \mathcal{K}^*,\ \ z_{ij} \in \mathbb{R}^n & \forall i \leq N, \forall j \leq d \\
& \displaystyle\sum_{j=1}^d \theta_j \Psi_j^*(z_{ij}/\theta_j) + \langle \theta, \Psi(\widehat{x}_i) \rangle + \langle \phi_{i1}, C\widehat{s}_i - d \rangle - \langle \gamma_i, H\widehat{s}_i + h \rangle \leq r_i + \tau & \forall i \leq N \\
& \displaystyle\sum_{j=1}^d z_{ij} = W^\top \gamma_i & \forall i \leq N \\
& \langle C\widehat{s}_i - d, \phi_{i2} \rangle \leq r_i & \forall i \leq N \\
& \left\| \begin{pmatrix} H^\top \gamma_i - C^\top \phi_{i1} \\ \theta \end{pmatrix} \right\|_* \leq \lambda, \quad \left\| \begin{pmatrix} C^\top \phi_{i2} \\ 0 \end{pmatrix} \right\|_* \leq \lambda & \forall i \leq N.
\end{array} \tag{35}$$

Note that Theorem A.1 remains valid if $(\widehat{s}_i, \widehat{x}_i) \notin \Xi$ for some $i \leq N$.

Proof of Theorem A.1 As in the proof of Theorem 5.2 one can show that the objective function of the inverse optimization problem (13) is equivalent to (19). In the remainder, we derive a safe conic approximation for the (intractable) subordinate worst-case expectation problem

$$\sup_{\mathbb{Q}\in\mathbb{B}_\varepsilon^1(\widehat{\mathbb{P}}_N)} \mathbb{E}^{\mathbb{Q}}\big[\max\{\ell_\theta(s,x)-\tau,0\}\big]. \tag{36}$$

To this end, note that the suboptimality loss $\ell_\theta(s,x)=\langle\theta,\Psi(x)\rangle-\min_{y\in\mathbb{X}(s)}\langle\theta,\Psi(y)\rangle$ depends on x only through $\Psi(x)$. This motivates us to define a lifted suboptimality loss $\ell_\theta^\Psi(s,\psi)=\langle\theta,\psi\rangle-\min_{y\in\mathbb{X}(s)}\langle\theta,\Psi(y)\rangle$, where ψ represents an element of the feature space $\mathbb{R}^d$, and the empirical distribution $\widehat{\mathbb{P}}_N^\Psi=\frac{1}{N}\sum_{i=1}^N\delta_{(\widehat{s}_i,\Psi(\widehat{x}_i))}$ of the signal-feature pairs. Moreover, we denote by $\mathbb{B}_\varepsilon^1(\widehat{\mathbb{P}}_N^\Psi)$ the 1-Wasserstein ball of all distributions on $\mathbb{S}\times\mathbb{R}^d$ that have a distance of at most ε from $\widehat{\mathbb{P}}_N^\Psi$. In the following we will show that the worst-case expectation

$$\sup_{\mathbb{Q}\in\mathbb{B}_\varepsilon^1(\widehat{\mathbb{P}}_N^\Psi)} \mathbb{E}^{\mathbb{Q}}\big[\max\{\ell_\theta(s,\psi)-\tau,0\}\big] \tag{37}$$

on the signal-feature space $\mathbb{S}\times\mathbb{R}^d$ provides a tractable upper bound on the worst-case expectation (36). By Definition 4.3, each distribution $\mathbb{Q}\in\mathbb{B}_\varepsilon^1(\widehat{\mathbb{P}}_N)$ corresponds to a transportation plan Π, that is, a joint distribution of two signal-response pairs (s,x) and (s',x') under which (s,x) has marginal distribution $\mathbb{Q}$ and (s',x') has marginal distribution $\widehat{\mathbb{P}}_N$. By the law of total probability, the transportation plan can be expressed as $\Pi=\frac{1}{N}\sum_{i=1}^N\delta_{(\widehat{s}_i,\widehat{x}_i)}\otimes\mathbb{Q}_i$, where $\mathbb{Q}_i$ denotes the conditional distribution of (s,x) given $(s',x')=(\widehat{s}_i,\widehat{x}_i)$, $i\le N$, see also [30, Theorem 4.2]. Thus, we the worst-case expectation (36) satisfies

$$\begin{aligned}
&\sup_{\mathbb{Q}\in\mathbb{B}_\varepsilon^1(\widehat{\mathbb{P}}_N)} \mathbb{E}^{\mathbb{Q}}\big[\max\{\ell_\theta(s,x)-\tau,0\}\big]\\
&=\sup_{\mathbb{Q}^i}\frac{1}{N}\sum_{i=1}^N\int_\Xi\max\{\ell_\theta(s,\Psi(x))-\tau,0\}\,\mathbb{Q}^i(\mathrm{d}s,\mathrm{d}x)\\
&\text{s.t.}\frac{1}{N}\sum_{i=1}^N\int_\Xi\|(s,x)-(\widehat{s}_i,\widehat{x}_i)\|\,\mathbb{Q}^i(\mathrm{d}s,\mathrm{d}x)\le\varepsilon\\
&\int_\Xi\mathbb{Q}^i(\mathrm{d}s,\mathrm{d}x)=1\qquad\forall i\le N\\
&\le\sup_{\mathbb{Q}^i}\frac{1}{N}\sum_{i=1}^N\int_\Xi\max\{\ell_\theta(s,\Psi(x))-\tau,0\}\,\mathbb{Q}^i(\mathrm{d}s,\mathrm{d}x)\\
&\text{s.t.}\frac{1}{N}\sum_{i=1}^N\int_\Xi\|(s,\Psi(x))-(\widehat{s}_i,\Psi(\widehat{x}_i))\|\,\mathbb{Q}^i(\mathrm{d}s,\mathrm{d}x)\le\varepsilon\\
&\int_\Xi\mathbb{Q}^i(\mathrm{d}s,\mathrm{d}x)=1\qquad\forall i\le N\\
&\le\sup_{\mathbb{Q}^i}\frac{1}{N}\sum_{i=1}^N\int_{\mathbb{S}\times\mathbb{R}^d}\max\{\ell_\theta(s,\psi)-\tau,0\}\,\mathbb{Q}^i(\mathrm{d}s,\mathrm{d}\psi)
\end{aligned}$$

$$\text{s.t.} \frac{1}{N}\sum_{i=1}^{N}\int_{\mathbb{S}\times\mathbb{R}^d} \|(s,\psi)-(\widehat{s}_i,\Psi(\widehat{x}_i))\|\,\mathbb{Q}^i(\mathrm{d}s,\mathrm{d}\psi)\le\varepsilon$$
$$\int_{\mathbb{S}\times\mathbb{R}^d}\mathbb{Q}^i(\mathrm{d}s,\mathrm{d}\psi)=1 \qquad \forall i\le N,$$

where the first inequality follows from (34), while the second inequality follows from relaxing the implicit condition that the signal-feature pair (s,ψ) must be supported on $\{(s,\Psi(x)) : (s,x)\in\Xi\}\subseteq\mathbb{S}\times\mathbb{R}^d$. Using a similar reasoning as before, the last expression is readily recognized as the worst-case expectation (37). Thus, (37) provides indeed an upper bound on (36). Duality arguments borrowed from [30, Theorem 4.2] imply that the infinite-dimensional linear program (37) admits a strong dual of the form

$$\begin{array}{ll}\inf\limits_{\lambda\ge 0,r_i} & \lambda\varepsilon+\dfrac{1}{N}\displaystyle\sum_{i=1}^{N} r_i\\ \text{s.t.} & \sup\limits_{(s,y)\in\Xi,\psi\in\mathbb{R}^d}\big\langle\theta,\psi-\Psi(y)\big\rangle-\tau-\lambda\|(s,\psi)-(\widehat{s}_i,\Psi(\widehat{x}_i))\|\le r_i \quad \forall i\le N\\ & \sup\limits_{s\in\mathbb{S},\psi\in\mathbb{R}^d} -\lambda\|(s,\psi)-(\widehat{s}_i,\Psi(\widehat{x}_i))\|\le r_i \qquad\qquad \forall i\le N.\end{array} \tag{38}$$

By using the definitions of $\mathbb{S}$ and $\mathbb{X}(s)$ put forth in Assumption 5.1, the i-th member of the first constraint group in (38) is satisfied if and only if the optimal value of the maximization problem

$$\begin{array}{ll}\sup\limits_{s,y,\psi} & \big\langle\theta,\psi-\Psi(y)\big\rangle-\tau-\lambda\|(s,\psi)-(\widehat{s}_i,\Psi(\widehat{x}_i))\|\\ \text{s.t.} & Cs\succeq_{\mathcal{C}} d,\ \ Wy\succeq_{\mathcal{K}} Hs+h\end{array} \tag{39}$$

does not exceed r_i. A tedious but routine calculation shows that the dual of (39) can be represented as

$$\begin{array}{ll}\inf\limits_{p_i,q_i,\gamma_i,\phi_{i1},z_{ij}} & \displaystyle\sum_{j=1}^{d}\theta_j\Psi_j^*(z_{ij}/\theta_j)-\tau+\big\langle\theta,\Psi(\widehat{x}_i)\big\rangle+\big\langle\phi_{i1},C\widehat{s}_i-d\big\rangle-\big\langle\gamma_i,H\widehat{s}_i+h\big\rangle\\ \text{s.t.} & \sum_{j=1}^{d} z_{ij}=W^\top\gamma_i\\ & \|(H^\top\gamma_i-C^\top\phi_{i1},\theta)\|_*\le\lambda\\ & \gamma_i\in\mathcal{K}^*,\ \phi_{i1}\in\mathcal{C}^*.\end{array} \tag{40}$$

Note that the perspective functions $\theta_j\Psi_j^*(z_{ij}/\theta_j)$ in the objective of (40) emerge from taking the conjugate of $\theta_j\Psi_j(y)$. Thus, for $\theta_j=0$ we must interpret $\theta_j\Psi_j^*(z_{ij}/\theta_j)$ as an indicator function in z_{ij} which vanishes for $z_{ij}=0$ and equals ∞ otherwise. By weak duality, (40) provides an upper bound on (39). We conclude that the i-th member of the first constraint group in (38) is satisfied whenever the dual problem (40) has a feasible solution whose objective value does not exceed r_i. A similar reasoning shows

that the i-th member of the second constraint group in (38) holds if and only if there exists $\phi_{i2} \in \mathcal{C}^*$ such that

$$\left\langle C\widehat{s}_i - d, \phi_{i2} \right\rangle \leq r_i \quad \text{and} \quad \left\| \begin{pmatrix} C^\intercal \phi_{i2} \\ 0 \end{pmatrix} \right\|_* \leq \lambda.$$

Thus, the worst-case expectation (36) is bounded above by the optimal value of the finite convex program

$$\begin{array}{lll}
\inf & \varepsilon\lambda + \dfrac{1}{N}\displaystyle\sum_{i=1}^{N} r_i & \\
\text{s.t.} & \lambda \in \mathbb{R}_+,\ r_i \in \mathbb{R},\ \phi_{i1}, \phi_{i2} \in \mathcal{C}^*,\ \gamma_i \in \mathcal{K}^* & \forall i \leq N \\
& \sum_{j=1}^{d} \theta_j \Psi_j^*(z_{ij}/\theta_j) + \left\langle \theta, \Psi(\widehat{x}_i) \right\rangle + \left\langle \phi_{i1}, C\widehat{s}_i - d \right\rangle - \left\langle \gamma_i, H\widehat{s}_i + h \right\rangle \leq r_i + \tau & \forall i \leq N \\
& \sum_{j=1}^{d} z_{ij} = W^\top \gamma_i & \forall i \leq N \\
& \left\langle C\widehat{s}_i - d, \phi_{i2} \right\rangle \leq r_i & \\
& \left\| \begin{pmatrix} H^\top \gamma_i - C^\top \phi_{i1} \\ \theta \end{pmatrix} \right\|_* \leq \lambda, \quad \left\| \begin{pmatrix} C^\top \phi_{i2} \\ 0 \end{pmatrix} \right\|_* \leq \lambda & \forall i \leq N.
\end{array}$$

The claim then follows by substituting this convex program into (19).

References

1. Ackerberg, D., Benkard, C.L., Berry, S., Pakes, A.: Econometric tools for analyzing market outcomes. In: Heckman, J., Leamer, E. (eds.) Handbook of Econometrics, pp. 4171–4276. Elsevier, Amsterdam (2007)
2. Ahmed, S., Guan, Y.: The inverse optimal value problem. Math. Program. A **102**, 91–110 (2005)
3. Ahuja, R.K., Orlin, J.B.: A faster algorithm for the inverse spanning tree problem. J. Algorithms **34**, 177–193 (2000)
4. Ahuja, R.K., Orlin, J.B.: Inverse optimization. Oper. Res. **49**, 771–783 (2001)
5. Anderson, B.D., Moore, J.B.: Optimal Control: Linear Quadratic Methods. Courier Corporation, North Chelmsford (2007)
6. Aswani, A., Shen, Z.-J.M., Siddiq, A.: Inverse optimization with noisy data. Preprint arXiv:1507.03266 (2015)
7. Bajari, P., Benkard, C.L., Levin, J.: Estimating dynamic models of imperfect competition. Technical Report, National Bureau of Economic Research (2004)
8. Ben-Ayed, O., Blair, C.E.: Computational difficulties of bilevel linear programming. Oper. Res. **38**, 556–560 (1990)
9. Ben-Tal, A., Ghaoui, El, Nemirovski, A.: Robust Optimization. Princeton University Press, Princeton (2009)
10. Bertsimas, D., Gupta, V., Paschalidis, I.C.: Inverse optimization: a new perspective on the Black–Litterman model. Oper. Res. **60**, 1389–1403 (2012)
11. Bertsimas, D., Gupta, V., Paschalidis, I.C.: Data-driven estimation in equilibrium using inverse optimization. Math. Program. **153**, 595–633 (2015)
12. Boissard, E.: Simple bounds for convergence of empirical and occupation measures in 1-Wasserstein distance. Electron. J. Probab. **16**, 2296–2333 (2011)
13. Bottasso, C.L., Prilutsky, B.I., Croce, A., Imberti, E., Sartirana, S.: A numerical procedure for inferring from experimental data the optimization cost functions using a multibody model of the neuromusculoskeletal system. Multibody Syst. Dyn. **16**, 123–154 (2006)
14. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge (2009)
15. Braess, D., Nagurney, A., Wakolbinger, T.: On a paradox of traffic planning. Transp. Sci. **39**, 446–450 (2005)

16. Bubeck, S.: Convex optimization: algorithms and complexity. Found. Trends Mach. Learn. **8**, 231–357 (2015)
17. Burton, D., Toint, P.L.: On an instance of the inverse shortest paths problem. Math. Program. **53**, 45–61 (1992)
18. Carr, S., Lovejoy, W.: The inverse newsvendor problem: choosing an optimal demand portfolio for capacitated resources. Manag. Sci. **46**, 912–927 (2000)
19. Chan, T.C., Craig, T., Lee, T., Sharpe, M.B.: Generalized inverse multiobjective optimization with application to cancer therapy. Oper. Res. **62**, 680–695 (2014)
20. Faragó, A., Szentesi, Á., Szviatovszki, B.: Inverse optimization in high-speed networks. Discrete Appl. Math. **129**, 83–98 (2003)
21. Fournier, N., Guillin, A.: On the rate of convergence in Wasserstein distance of the empirical measure. Probab. Theory Relat. Fields **162**, 707–738 (2014)
22. Friedman, J., Hastie, T., Tibshirani, R.: The Elements of Statistical Learning. Springer, Berlin (2001)
23. Harker, P., Pang, J.: Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications. Math. Program. **48**, 161–220 (1990)
24. Heuberger, C.: Inverse combinatorial optimization: a survey on problems, methods, and results. J. Comb. Optim. **8**, 329–361 (2004)
25. Hochbaum, D.S.: Efficient algorithms for the inverse spanning-tree problem. Oper. Res. **51**, 785–797 (2003)
26. Iyengar, G., Kang, W.: Inverse conic programming with applications. Oper. Res. Lett. **33**, 319–330 (2005)
27. Keshavarz, A., Wang, Y., Boyd, S.: Imputing a convex objective function. In: IEEE International Symposium on Intelligent Control, pp. 613–619 (2011)
28. Lofberg, J.: YALMIP: a toolbox for modeling and optimization in MATLAB. In: IEEE International Symposium on Computer Aided Control Systems Design, pp. 284 –289 (2004)
29. Markowitz, H .M.: Portfolio Selection: Efficient Diversification of Investments. Yale University Press, New Haven (1968)
30. Mohajerin Esfahani, P., Kuhn, D.: Data-driven distributionally robust optimization using the Wasserstein metric: performance guarantees and tractable reformulations. Math. Program. (2017). https://doi.org/10.1007/s10107-017-1172-1
31. Murty, K.G., Kabadi, S.N.: Some NP-complete problems in quadratic and nonlinear programming. Math. Program. **39**, 117–129 (1987)
32. Neumann-Denzau, G., Behrens, J.: Inversion of seismic data using tomographical reconstruction techniques for investigations of laterally inhomogeneous media. Geophys. J. Int. **79**, 305–315 (1984)
33. Rockafellar, R.T., Uryasev, S.: Optimization of conditional value-at-risk. J. Risk **2**, 21–42 (2000)
34. Saez-Gallego, J., Morales, J.M., Zugno, M., Madsen, H.: A data-driven bidding model for a cluster of price-responsive consumers of electricity. IEEE Trans. Power Syst. **31**, 5001–5011 (2016)
35. Schaefer, A.J.: Inverse integer programming. Optim. Lett. **3**, 483–489 (2009)
36. Shafieezadeh-Abadeh, S., Kuhn, D., Mohajerin Esfahani, P.: Regularization via mass transportation. Preprint arXiv:1710.10016 (2017)
37. Shafieezadeh-Abadeh, S., Esfahani, P.M., Kuhn, D.: Distributionally robust logistic regression. In: Cortes, C., Lawrence, N.D., Lee, D.D., Sugiyama, M., Garnett, R. (eds.) Advances in Neural Information Processing Systems, vol. 28, pp. 1576–1584. Curran Associates, Inc. (2015)
38. Sion, M.: On general minimax theorems. Pac. J. Math. **8**, 171–176 (1958)
39. Terekhov, A.V., Pesin, Y.B., Niu, X., Latash, M.L., Zatsiorsky, V.M.: An analytical approach to the problem of inverse optimization with additive objective functions: an application to human prehension. J. Math. Biol. **61**, 423–453 (2010)
40. Troutt, M.D., Pang, W.-K., Hou, S.-H.: Behavioral estimation of mathematical programming objective function coefficients. Manag. Sci. **52**, 422–434 (2006)
41. Wang, L.: Cutting plane algorithms for the inverse mixed integer linear programming problem. Oper. Res. Lett. **37**, 114–116 (2009)
42. Woodhouse, J.H., Dziewonski, A.M.: Mapping the upper mantle: three-dimensional modeling of earth structure by inversion of seismic waveforms. J. Geophys. Res. **89**, 5953–5986 (1984)
43. Zhang, J., Xu, C.: Inverse optimization for linearly constrained convex separable programming problems. Eur. J. Oper. Res. **200**, 671–679 (2010)