Probabilistic Constraints for Inverse Problems

Elsa Carvalho, Jorge Cruz, and Pedro Barahona

Centro de Inteligência Artificial, Universidade Nova de Lisboa, Portugal
elsac@uma.pt, {jc,pb}@di.fct.unl.pt

Summary. The authors previous work on probabilistic constraint reasoning assumes the uncertainty of numerical variables within given bounds, characterized by *a priori* probability distributions. It propagates such knowledge through a network of constraints, reducing the uncertainty and providing *a posteriori* probability distributions. An inverse problem aims at estimating parameters from observed data, based on some underlying theory about a system behavior. This paper describes how nonlinear inverse problems can be cast into the probabilistic constraint framework, highlighting its ability to deal with all the uncertainty aspects of such problems.

1 Introduction

Many problems of practical interest can be formulated as nonlinear inverse problems. Such problems aim at finding the parameters of a model, given by systems of equations, from noisy data. Classical approaches for these problems are based on nonlinear regression methods, which search for the model parameter values that best-fit a given criterion [9]. Best-fit approaches, often based on local search methods, provide a non reliable single solution which may not be enough to the adequate characterization of the parameters.

Other stochastic approaches [16] associate a probabilistic model to the problem, from which is possible to obtain any sort of statistical information on the model parameters. These approaches typically rely on extensive random sampling to characterize the parameter space. However, even after intensive computations, no definitive conclusion can be drawn with these approaches, because a significant subset of the parameter space may have been missed.

In contrast, bounded-error approaches [5, 4] aim at characterizing the set of all solutions consistent with the uncertainty on the parameters, the model and the data. This is achieved through constraint reasoning, where initial intervals, representing the uncertainty on parameter values, are safely narrowed by reliable interval methods. Nevertheless, this approach has a major pitfall as it considers the same likelihood for all values in the intervals. In the interval computations context a combination of intervals with probabilistic uncertainty was proposed [6]. However, its dependence on a forward evaluation algorithm makes it unsuitable for inverse problems.

V.-N. Huynh et al. (Eds.): Interval/Probabilistic Uncertainty, ASC 46, pp. 115–128, 2008.
springerlink.com

The authors previous proposal of a probabilistic constraint framework [2] aims at computing an *a posteriori* distribution from an *a priori* distribution accordingly to safe reasoning on a continuous constraint model. In this paper we argue that such framework may constitute an attractive alternative approach to inverse problems, bridging the gap between pure probabilistic reasoning and pure safe reasoning.

The paper is organized as follows. In section 2 inverse problems are introduced. In sections 3 and 4 the constraint programming framework and the basic ideas of probabilistic reasoning are presented, together with their existing approaches to inverse problems. Next probabilistic interval computations are briefly presented. Section 6 describes the authors probabilistic constraint framework and highlights its ability to deal with all the uncertainty aspects of inverse problems. Finally, conclusions and future work are discussed.

2 Inverse Problems

A mathematical model describes a system by a set of variables and equations that establish relationships between them. In the context of inverse problems, the variables are divided into model parameters, whose values completely characterize the system, and observable parameters, which can be measured. The model is typically a forward model, defining a mapping from the model parameters to the observable parameters. It allows predicting the results of measurements based on the model parameters. An inverse problem is the task of obtaining values for the model parameters from the observed data.

The forward mapping, resulting from some theory about the system behavior, is commonly represented as a vector function $\boldsymbol{f}$ from the parameter space $\mathbf{m}$ (model parameters) to the data space $\mathbf{d}$ (observable parameters):

$$\mathbf{d} = \boldsymbol{f}(\mathbf{m}) \tag{1}$$

Such relation may be represented explicitly by an analytical formula or implicitly by a complex system of equations or some special purpose algorithm.

Nonlinearity and uncertainty play a major role in modeling the behavior of most real systems. In inverse problems the main sources of uncertainty are model approximations and measurement errors. Given uncertainty, an inverse problem may have no exact solutions, since usually there are no model parameter values capable of predicting exactly all the observed data. However, since the model equations are often highly nonlinear, uncertainty may be dramatically magnified, and an arbitrarily small change in the data may induce an arbitrarily large change in the values of the model parameters.

For example, consider the data summarized in Table 1 based on the USA census over the years 1790 (normalized to 0) to 1910 with a 10 year period.

Assuming that an exponential growth is an acceptable model for the population growth, the forward mapping would be defined by the following set of equations (one for each pair $\langle t_i, d_i \rangle$):

Table 1. US Population (in millions) over the years 1790 (0) to 1910 (120)

t_i	0	10	20	30	40	50	60	70	80	90	100	110	120
d_i	3.9	5.3	7.2	9.6	12.9	17.1	23.2	31.4	39.8	50.2	62.9	76.0	92.0

$$d_i = m_0 e^{m_1 t_i} \tag{2}$$

where m_0 and m_1 are the model parameters whose values must be estimated from the observed data. This is an example of an inverse problem with a simple nonlinear forward model but with no possible value combination for the model parameters satisfying the observed data set.

This problem is classically handled as a curve fitting problem [9]. Such approaches for nonlinear inverse problems are based on nonlinear regression methods which search for the model parameter values that best-fit a given criterion. For instance, the (weighted) least squares criterion minimizes a (weighted) quadratic norm of the difference between the vector of observed data and the vector of model predictions.

In the above example the weighted least squares criterion would be the minimization of the expression:

$$\sum_i \left(\frac{d_i - m_0 e^{m_1 t_i}}{\sigma_i} \right)^2 \tag{3}$$

where σ_i is the weight associated with the error of the i^{th} observation.

The minimization criteria (such as the weighted least squares criterion) are justified by the hypothesis that all problem uncertainties may be modeled using some well behaved distributions (such as Gaussians) eventually with specific parameter values for the different observations (the σ_i values). This is the case for some linear or weakly nonlinear inverse problems, where efficient computational techniques exist to solve them as curve fitting problems [9].

In generic nonlinear inverse problems, where no explicit formula can be provided for obtaining the best-fit values, minimization is often performed through local search algorithms. However, the search method may stop at a local minimum with no guarantees on the complete search space. Moreover, in most problems, a single best-fit solution may not be enough. Since other solutions could also be quite satisfactory with respect to the adopted criterion, the uncertainty around possible solutions should also be characterized. Analytic techniques can only be used for this purpose relying on some special assumptions about the model parameter distributions (for instance, assuming a single maximum). However, if the problem is highly nonlinear such assumptions do not provide realistic approximations for the uncertainty.

3 Continuous Constraint Satisfaction Problems

A Constraint Satisfaction Problem (CSP) [10] is defined by a triple (X;D;C) where X is a set of variables, each with an associated domain of possible

values in D, and C is a set of constraints on subsets of the variables. A constraint specifies which values from the domains of its variables are compatible. A solution to the CSP is an assignment of values to all its variables, which satisfies all the constraints. In continuous CSPs (CCSPs) [8, 1, 15] variable domains are continuous real intervals and constraints are equalities and inequalities. The space of possibilities is represented by boxes, i.e., the Cartesian product of real intervals. The CCSP framework is powerful enough to model a wide range of problems. In particular, engineering systems with components described as sets of continuous valued variables and relations defined by numerical equalities or inequalities, eventually with uncertain parameters. Continuous constraint reasoning eliminates value combinations from the initial search space (the Cartesian product of the initial domains), without loosing solutions. It combines pruning and branching steps until a stopping criterion is satisfied.

The pruning of the variable domains is based on constraint propagation. The main idea is to use the partial information expressed by a constraint to eliminate some incompatible values from the domain of its variables. Once the domain of a variable is reduced, this information is propagated to all constraints with that variable in their scopes. The process terminates when the domains cannot be further reduced by any constraint. Safe narrowing functions (mappings between boxes) are associated with constraints, to eliminate incompatible value combinations. Efficient methods from interval analysis (e.g. the interval Newton [11]) are often used to implement efficient narrowing functions which are correct (do not eliminate solutions) and contracting (the box obtained is smaller or equal than the original one).

Constraint propagation is a local consistency algorithm for pruning the variable domains, which is often insufficient to support safe decisions. To obtain better pruning, it is necessary to split the boxes and reapply constraint propagation to each sub-box. Such branch and prune process enforces a stronger, non local, consistency criterion. Several consistency criteria have been proposed [8, 1, 3], with distinct trade-offs between efficiency and pruning.

3.1 Constraint Approach to Inverse Problems

In the classical CCSP framework, the uncertainty associated with the problem is modeled by using intervals to represent the domains of the variables. Constraint reasoning reduces uncertainty, by reshaping the search space to become a safe approximation of the solution space. Such framework, with its efficient techniques to deal with nonlinear constraints and its safe uncertainty narrowing capabilities, is naturally appealing to handle nonlinear inverse problems beyond the classical best fitting methods.

The application of CCSPs in the context of inverse problems is known as bounded-error estimation or set membership estimation [5, 4]. The idea is to replace the search for a single best-fit solution (a parameter value combination) with the characterization of the set of all solutions consistent with the forward model, the uncertainty on the parameters and on the observations.

In its simplest form, bounded-error estimation assumes initial intervals to each problem variable, either a model or an observable parameter, and solve the CCSP with the set of constraints representing the forward model. Such strategy assumes prior knowledge on the acceptable parameter ranges as well as on the uncertainty of (difference between) predicted and observed data.

From the safe approximation of the solution space, a projection on the set of model parameters (or any subset of it) provides insight on the remaining uncertainty about their possible value combinations. In practice, since the narrowed ranges of the observable parameters are not essential, the variables that represent them may be replaced by their respective initial intervals.

Applying bounded-error estimation to the inverse problem described in section 2, it can be reformulated as a CCSP with the following set of constraints (one for each pair $\langle t_i, d_i \rangle$):

$$[d_i - \delta_i, d_i + \delta_i] \ni m_0 e^{m_1 t_i} \tag{4}$$

where δ_i is an acceptable difference between the i^{th} observation and the respective predicted value, and m_0 and m_1 are the only variables of the constraint model. The initial ranges for these variables should be provided within reasonable bounds (I_0 and I_1) and represent the parameter uncertainty that will be reduced through constraint reasoning. Figure 1 shows the approximation of the solution space that is computed with $I_0 = [0, 100]$, $I_1 = [0.01, 0.1]$ and $\delta_i = 3$ for all observations presented in Table 1. From the figure, it is clear which combinations of the model parameter values are consistent with the initial uncertainty assumptions, the forward model and the observations.

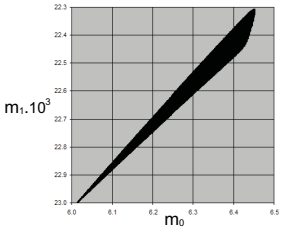

Fig. 1. Approximation of the CCSP solution space

The formulation of an inverse problem as a CCSP may easily accommodate additional requirements, in the form of constraints, which are more difficult to enforce in classical approaches. Moreover, the generality of this approach allows its application to inverse problems whose forward model is not defined by an explicit analytical formula but rather by a complex set of relations.

However, in many cases, safe reasoning is useless, intervals are often very wide, and subsequent constraint propagation is not able to narrow them. In fact, an uncertain value may range over a wide interval but a much narrower interval may include the most likely values. In some problems, the plausibility distribution of values within the bounds of an uncertain parameter is known. For

instance, uncertainty due to measuring errors may be naturally associated with an error distribution. However, the traditional CCSP framework cannot accommodate such information and thus, for each variable, all values in its domain are considered equally plausible.

4 Probabilistic Reasoning

Probability provides a classical model for dealing with uncertainty. The basic element of probability theory is the random variable, which plays a similar role to that of the CSP variables. Each random variable has a domain where it can assume values. In particular, continuous random variables assume real values. A possible world, or atomic event, is an assignment of values to all the variables of the model. An event is a set of possible worlds. The complete set of all possible worlds in the model is the sample space. If all the random variables are continuous, the sample space is the hyperspace obtained by the Cartesian product of the variable domains, and the possible worlds and events are, respectively, points and regions from such hyperspace.

Probability measures may be associated with events. In the continuous case, an assignment of a probability to a point, is representative of the likelihood in its neighborhood. A probabilistic model is an encoding of probabilistic information, allowing to compute the probability of any event, in accordance with the axioms of probability. The usual method for specifying a probabilistic model assumes, either explicitly or implicitly, a full joint probability distribution, which assigns a probability measure to each possible world.

Probabilistic reasoning aims at incorporating new information, known as evidence, by updating an *a priori* probability into an *a posteriori* probability given the evidence. The *a priori* probability is a description of what is known in the absence of the evidence. For incorporating this evidence, conditioning is used. Conditional probability $P(A|B)$ is the probability of some event A, given the occurrence of some other event B. The *a posteriori* probability is the conditional probability when the relevant evidence is taken into account.

Probabilistic graphical models [7] (Markov networks and Bayesian networks [13]) provide a powerful framework for efficient probabilistic reasoning. The idea is to use a probabilistic network that captures the structural properties of the probabilistic model (such as conditional independence) and defines an implicit full joint probability distribution. Given new evidence (information about some nodes), belief propagation [13] is one of the most efficient inference algorithms to compute *a posteriori* probabilities for all the non-evidence nodes in the network. However, such approaches, which require the full specification of a conditional probability at each node of the network, are often inadequate for continuous nonlinear problems.

4.1 Probabilistic Approach to Inverse Problems

Inverse problems are often handled by probabilistic approaches that associate an explicit probabilistic model to the problem [16]. Prior information on the model

parameters is represented by a probability distribution, which is transformed into an *a posteriori* probability distribution, by incorporating a forward theory (relating the model parameters to the observable parameters) and the actual result of the observations (with their uncertainties).

On these approaches all the information related to inverse problems is described using probability densities. The *a posteriori* probability density of the model parameters $\sigma_M(\mathbf{m})$ can be computed from these distributions. For example, assuming linear data and model spaces this probability density is:

$$\sigma_M(\mathbf{m}) = k\rho_M(\mathbf{m}) \int_D \rho_D(\mathbf{d})\Theta(\mathbf{d}, \mathbf{m})d\mathbf{d} \tag{5}$$

where k is a normalization constant and $\rho_D(\mathbf{d})$, $\rho_M(\mathbf{m})$ and $\Theta(\mathbf{d}, \mathbf{m})$ are the probability densities for the observable parameters, the model parameters and the underlying theory.

$\sigma_M(\mathbf{m})$ provides an explicit definition of a full joint probability distribution on the model parameter space, from which is possible to obtain any sort of statistical information on the model parameters compatible with the *a priori* uncertainty, the theoretical information and the experimental results. In the particular case where $\sigma_M(\mathbf{m})$ is identically null, some incompatible assumptions were surely made, indicating that uncertainty has been underestimated.

Only in very simple cases analytic techniques can be used to characterize $\sigma_M(\mathbf{m})$. In general, it is necessary to perform an extensive exploration of the model space. When it is small, a systematic exploration may be achieved, computing $\sigma_M(\mathbf{m})$ at every point of a grid defined over the complete space. Usually, such exploration cannot be systematic since too many points would have to be evaluated, so it is replaced by a random (Monte Carlo) exploration. For a more detailed description of these methods see [16, 12].

Nevertheless, only a number of discrete points of the continuous model space is analyzed and the results must be extrapolated to characterize the overall uncertainty. Such approach is highly dependent on the exploration length which, to provide better uncertainty characterizations, need to be reinforced in highly nonlinear problems. Moreover, contrary to constraint reasoning approaches, these probabilistic techniques cannot prune the search space based on model information. Consequently the entire space is considered for exploration, independently of its *a posteriori* probability distribution, which can have null values for inconsistent subregions.

5 Probabilistic Interval Computations

A combination of probabilistic and interval representations of uncertainty appears in [6]. This approach uses interval domains to represent the ranges of possible values and allows the incorporation of extra information about their probabilities. Such framework uses interval computations instead of the broader paradigm of continuous constraint reasoning. It is specially suited for data

processing problems, where an estimate for a quantity y must be computed, applying a known deterministic algorithm f to the available estimates for other quantities $x_1, \ldots, x_n$.

$$y = f(x_1, \ldots, x_n) \tag{6}$$

Data processing problems arise naturally when the estimates for the quantities $x_1, \ldots, x_n$ come from direct measurements obtained by physical instruments which provide upper bounds for the measurement errors, together with information about the error probability distributions. Probabilistic interval computations provide an alternative approach when only partial information about error distributions is available and standard statistical approaches [14] cannot be applied. It makes use of such partial information in the estimation of y. The idea is to maintain intervals to represent possible values of variables as well as possible values of parametric descriptors of their distributions (e.g., expected values). During interval computations such intervals are maintained consistent by a step by step evaluation process that extends basic interval arithmetic operations (see [6] for details).

Contrary to constraint approaches, that are based on undirected relations, this approach is highly dependent on the availability of a directed algorithm f to compute the pretended information for y from the $x_1, \ldots, x_n$ estimates. Clearly this is not the case for inverse problems where the model parameters are not outputs but rather inputs of the forward model. So, to apply this approach, it would be necessary to find a suitable analytical solution with respect to each model parameter.

6 Probabilistic Constraint Reasoning

In [2] the authors proposed the Probabilistic Continuous Constraint Satisfaction Problem (PCCSP) as an extension of a CCSP. A PCCSP is defined by (X;D;F;C), where X is a set of continuous random variables, each with an associated interval domain of possible values in D, distributed accordingly to the corresponding probability density function (p.d.f.) in F, and C is a set of constraints on subsets of the variables. Given a point in the domain of a random variable, its p.d.f. is representative of the *a priori* probability in its neighborhood, without considering the relations between the variables. It is assumed that all relevant relations between variables are expressed by the constraints of the model. Thus, when the constraints are not accounted for, the variables are independent.

The initial search space represents a probability space, characterized by a joint p.d.f. which, due to the independence assumption, is implicitly defined by the product of the individual p.d.f.s of the random variables. In the process of reducing uncertainty, there is a combination of continuous constraint reasoning and probabilistic reasoning. While the first reduces uncertainty by *reshaping* the search space, the second *redefines* the search space *a priori* probability distribution by computing an *a posteriori* distribution, based on the constraint reasoning outcome.

The constraints are the new information that is incorporated in the probabilistic model. The solution space is the event containing all possible worlds that satisfy the constraints. Through constraint reasoning an approximation (enclosure) of the solution space is obtained. Therefore the *a posteriori* probability is computed as a conditional probability, given the evidence represented by the approximation of the solution space. This probability is calculated by the conditional probability rule $P(A|B) = P(A \cap B)/P(B)$. The probability of region A given the evidence, is the probability of the subregion of A contained in the approximation of the solution space, divided by a normalizing factor.

The quality of the solution space approximation depends on the consistency and stopping criteria used in the constraint reasoning process. Regions of the search space that were not pruned during constraint reasoning may contain solutions, although there is no guaranty that they do. In fact, there is no knowledge why such regions are maintained. Was it due to lack of further exploration of this regions or did they contain solutions? Normally, the process of constraint reasoning, leads to non uniform sizes of the boxes that represent the solution space approximation. Nevertheless, for reasoning with probabilistic information, some kind of fairness in the exploration of the search space must be guaranteed, so that the obtained *a posteriori* distribution is not biased by heterogeneous search.

In the PCCSP, the stopping criterion is based on the maximum width ε_1 allowed for the intervals that constitute a box. A box is no further explored when all its intervals are smaller or equal to ε_1. When all the boxes meet this criterion the search stops. The stopping criterion assures some uniformity of the solution space approximation. However, due to the consistency enforcement narrowing capability, which differs between distinct regions of the search space, some heterogeneity is still present.

To maintain a generic non parametric representation of the *a posteriori* marginal p.d.f.s, some kind of discretization must be assumed. This is achieved by considering an ε_2-hypergrid, i.e., a grid where the dimension is the number n of variables in the PCCSP, and each grid unit (hypercube) has width ε_2 in all dimensions. The hypergrid allows to transform the non uniform solution space approximation, resulting from constraint reasoning, in a uniform one, providing a fair computation of the marginal p.d.f.s. The transformation is achieved by overlaying the hypergrid upon the solution space approximation, enforcing a *snap to grid* to this region. The new approximation is the set of grid hypercubes that intersect with the original approximation, producing a set of uniform boxes. Figure 2 illustrates the described reasoning process.

Once obtained the solution space approximation as a set SS of ε_2-hypergrid boxes, algorithm 1 calculates and returns the marginal *a posteriori* p.d.f. of a set of m variables (where $m \leq n$), discretized accordingly to the ε_2-hypergrid. For that purpose the algorithm maintains a m-dimensional matrix M, where each dimension corresponds to a variable. In the algorithm, $H_{box}[i]$ and $G_{box}[i]$ are the i^{th} intervals of the boxes (Cartesian product of n intervals). Given two intervals $I_1 = [l_1, r_1]$ and $I_2 = [l_2, r_2]$, the union hull $I_1 \uplus I_2$ is the interval

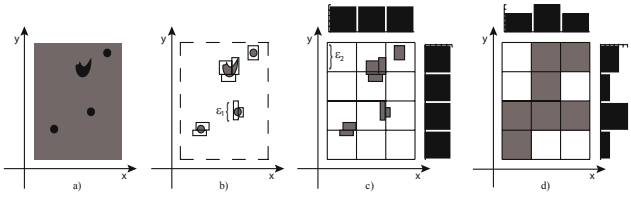

Fig. 2. Process of probabilistic constraint reasoning. (a) Initial search space and solution space; (b) Solution space approximation; (c) Hypergrid and *a priori* marginal p.d.f.s; (d) Snap to grid and *a posteriori* marginal p.d.f.s.

$[min(l_1, l_2), max(r_1, r_2)]$. H_{box} is a box where each interval is the union hull of the respective intervals of all the boxes in SS, i.e. is the smallest box enclosing all the boxes in SS (line 2). The length of each dimension is the number of ε_2 segments in which the corresponding variable domain can be divided (line 3). Each matrix cell thus obtained is initialized to zero (line 4) and its probability value is computed by summing up the contribution of all hypercubes (boxes) that are aligned with that cell (line 5-9), normalized by the sum of all hypercube contributions (lines 1, 10, 12). Due to the independence assumption, the contribution of an hypercube is the product of each variable contribution, i.e., the integral of its *a priori* p.d.f. (f_i), within the respective box interval (line 8).

Algorithm 1. Calculates marginal *a posteriori* p.d.f.

function $marginalAPosterioriPDF(SS, \varepsilon, m)$

1: $accum \leftarrow 0$
2: $H_{box} \leftarrow \uplus SS$
3: $\forall_{1 \leq i \leq m} \quad l_i \leftarrow H_{box}[i].width/\varepsilon$
4: $\forall_{1 \leq i_1 < l_1} \ldots \forall_{1 \leq i_m < l_m} \quad M[i_1] \ldots [i_m] \leftarrow 0$
5: **while** $SS \neq \oslash$ **do**
6: $\quad G_{box} \leftarrow remove(SS)$
7: $\quad \forall_{1 \leq i \leq m} \quad j_i \leftarrow (G_{box}[i].left - H_{box}[i].left)/\varepsilon$
8: $\quad p_{box} = \prod_{i=1}^{n} \int_{G_{box}[i].left}^{G_{box}[i].right} f_i(x_i)dx_i$
9: $\quad M[j_1] \ldots [j_m] \leftarrow M[j_1] \ldots [j_m] + p_{box}$
10: $\quad accum \leftarrow accum + p_{box}$
11: **end while**
12: $\forall_{1 \leq i_1 < n_1} \ldots \forall_{1 \leq i_m < n_m} \quad M[i_1] \ldots [i_m] \leftarrow M[i_1] \ldots [i_m]/accum$
13: **return** M

6.1 Probabilistic Constraint Approach to Inverse Problems

The application of PCCSPs in the context of inverse problems, based on bounded-error estimation, assumes both prior knowledge on the acceptable parameter ranges and on the uncertainty about the difference between predicted

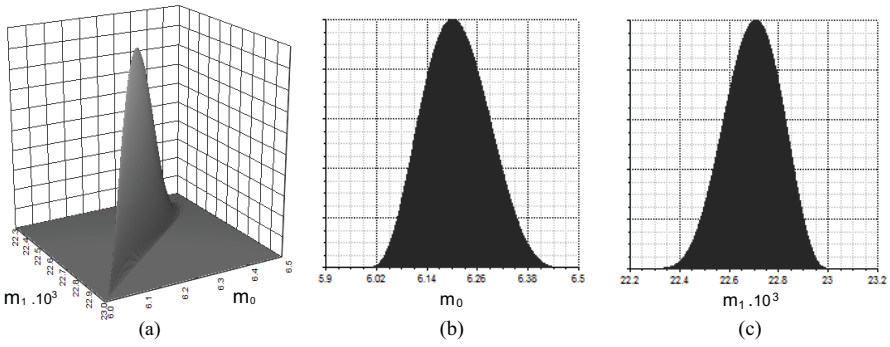

Fig. 3. Exponential model. (a) Joint p.d.f.; (b)(c) marginal p.d.f.s.

and observed data. This knowledge is expressed, respectively, by intervals and explicit *a priori* probability distributions within such intervals. If prior information is unavailable uniform distributions are considered.

When solving the PCCSP, where the set of constraints represents the forward model, a safe approximation of the solution space is obtained, and a projection on the set of model parameters (or any subset of it) provides insight on the *a posteriori* distribution of the resulting narrowed ranges.

In this approach, data parameters cannot be replaced by their respective initial intervals because it is necessary to keep track of their p.d.f.s for computing the *a posteriori* distributions. However, as long as the model parameters are the only shared variables between constraints, the contribution of each constraint, on the model parameters *a posteriori* distribution, may be independently computed and incrementally combined.

Consider again the inverse problem presented in section 2. Suppose that besides accepting a difference δ_i between the i^{th} observation and the respective predicted value, a p.d.f. $\rho_i(d_i)$ is associated to the acceptable interval representing the prior information on such difference. The initial ranges and respective p.d.f.s (possibly uniform distributions) must also be provided for the model parameters m_0 and m_1, characterizing the full *a priori* joint p.d.f.. In this case, the inverse problem may be reformulated as a PCCSP with the following set of constraints (one for each pair $\langle t_i, d_i \rangle$):

$$d_i = m_0 e^{m_1 t_i} \tag{7}$$

where m_0, m_1 and d_0, ... , d_{12} are the variables of the constraint model. The *a posteriori* distribution of the model parameters is computed by solving the PCCSP (projecting the results with respect to m_0 and m_1).

Figure 3 shows the *a posteriori* distribution of the model parameters that is computed with the initial ranges defined in subsection 3.1 and assuming *a priori* uniform distributions for the model parameters and triangular distributions (centered in d_i) for the observable parameters. Besides identifying which value combinations of m_0 and m_1 are consistent, figure 3(a) illustrates its joint

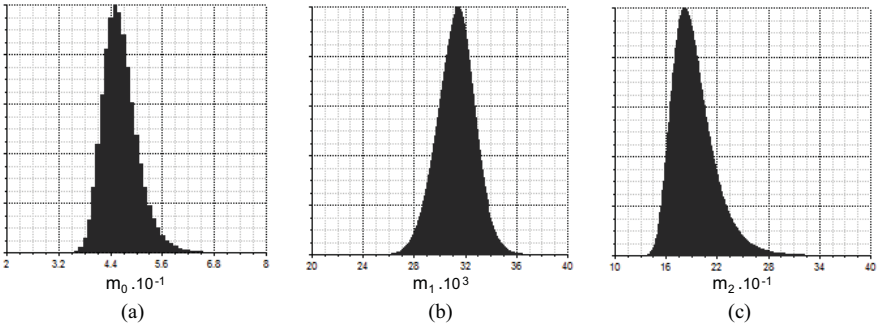

Fig. 4. Logistic model. (a)(b)(c) marginal p.d.f.s.

probability distribution, allowing to identify regions of maximum likelihood. Figures 3(b) and 3(c) are projections on m_0 and m_1 showing the *a posteriori* probability computed for each of the model parameters.

If a logistic model is considered, instead of an exponential model for the population growth, its reformulation as a PCCSP should keep the observable variables but satisfy a new set of constraints (one for each pair $\langle t_i, d_i \rangle$):

$$d_i = \frac{m_2}{1 + m_0 e^{-m_1 t_i}} \tag{8}$$

where m_0, m_1 and m_2 are the variables of the constraint model representing the model parameters. Figure 4 presents the marginal *a posteriori* distributions for each model parameter computed from the joint p.d.f. (with equal assumptions on the observed parameters uncertainty and with initially uniformly distributed $I_0 = [10, 100]$, $I_1 = [0.02, 0.05]$ and $I_2 = [100, 400]$).

The PCCSP associated with a given inverse problem can be easily extended to make predictions on the outcomes of new measurements. For this purpose a new constraint for each new measurement should be included in the model. Such constraints, similar to the other constraints representing the forward model, should include new unknown observable parameters (initially unbounded and uniformly distributed). *A posteriori* distributions for these new variables can be computed by solving the PCCSP and projecting the results with respect to each of them. Figure 5 illustrates the predictions for the population size in 1920 ($t_i = 130$) in the previous problem with both, the exponential model (figure 5(a)) and the logistic model (figure 5(b)). Note that the real observed value for the population size in 1920 was 106.0 (not shown in table 1) which is in accordance with the predictions of the logistic model, but outside the bounds predicted by the exponential model.

An insight about the quality of a particular model for a specific inverse problem may be achieved by analyzing the maximum likelihood regions. The obtained *a posteriori* marginal p.d.f. for the model parameters provides valuable information for inspecting the quality of a particular model. Not only it allows easy identification of maximum likelihood regions as peaks of such p.d.f., but also displays the complete shape of the uncertainty dispersion showing, for instance, if it is unimodal.

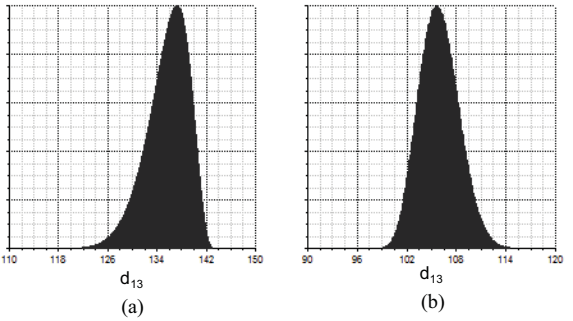

Fig. 5. Expected US population in 1920. (a) Exponential and (b) logistic models.

In the presented example, given the unimodality of the *a posteriori* p.d.f.s for both models, a quantitative measure of their quality may be obtained by evaluating any numerical best-fit criterion (see section 2) at their maximum likelihood points. The boxes that enclose such points for the exponential and the logistic models are, respectively, $\langle[6.159, 6.160], [0.022770, 0.022771]\rangle$ and $\langle[45, 46], [0.0318, 0.0319], [181, 182]\rangle$. The least squares criterion (formula 3 with $\sigma_i = 1$) evaluated at this boxes results, respectively, in $I_1 = [24.6, 18341.6]$ and $I_2 = [0.1, 11.8]$. Since the maximum likelihood points are included in those boxes and any value of I_2 is smaller than any value of I_1, according to the chosen criterion, the logistic model is a better representation for the population growth than the exponential model.

7 Conclusions and Future Work

This paper describes how inverse problems can be cast into the probabilistic continuous constraint framework. The approach introduces new expressive power for modeling the underlying theory about the system behavior and produces appealing graphical results for representing the uncertainty on model parameters and predictions on measurement outcomes. However, it seems to be more adequate to handle inverse problems with a reduced number of parameters. This is particularly true when the model is highly nonlinear, in which case, a smaller granularity is required for pruning the search space. To address the scalability of the approach further experimentation must be done on more realistic inverse problems. Furthermore, we intend to develop an interactive prototype, to improve usability and fully explore the framework capabilities.

References

1. Benhamou, F., McAllester, D., van Hentenryck, P.: Clp(intervals) revisited. In: ISLP, pp. 124–138. MIT Press, Cambridge (1994)
2. Carvalho, E., Cruz, J., Barahona, P.: Probabilistic reasoning with continuous constraints. In: AIP Conference Proceedings, vol. 936(1), pp. 105–108 (2007)

3. Cruz, J., Barahona, P.: Maintaining global hull consistency w/ local search for cc-sps. In: Bliek, C., Jermann, C., Neumaier, A. (eds.) COCOS 2002. LNCS, vol. 2861, pp. 178–193. Springer, Heidelberg (2003)
4. Granvilliers, L., Cruz, J., Barahona, P.: Parameter estimation using interval computations. SIAM J. Scientific Computing 26(2), 591–612 (2004)
5. Jaulin, L., Walter, E.: Set inversion via interval analysis for nonlinear bounded-error estimation. Automatica 29(4), 1053–1064 (1993)
6. Kreinovich, V.: Probabilities, intervals, what next? optimization problems related to extension of interval computations to situations with partial information about probabilities. Journal of Global Optimization 29, 265–280 (2004)
7. Lauritzen, S.L.: Graphical Models. Oxford Statistical Science Series, vol. 17. Clarendon Press, Oxford (1996)
8. Lhomme, O.: Consistency techniques for numeric csps. In: Proc. of the 13th IJCAI, pp. 232–238 (1993)
9. Moler, C.B.: Numerical Computing with Matlab. SIAM (2004)
10. Montanari, U.: Networks of constraints: Fundamental properties and application to picture processing. Information Sciences 7(2), 95–132 (1974)
11. Moore, R.: Interval Analysis. Prentice-Hall, Englewood Cliffs (1966)
12. Mosegaard, K.: Resolution analysis of general inverse problems through inverse monte carlo sampling. Inverse Problems 14(3), 405–426 (1998)
13. Pearl, J.: Probabilistic reasoning in intelligent systems: Networks of plausible inference. Morgan Kaufmann, San Francisco (1988)
14. Rabinovich, S.: Measurement Errors and Uncertainties: Theory and Practice. Springer, Heidelberg (2005)
15. Sam-Haroud, D., Faltings, B.: Consistency techniques for continuous constraints. Constraints 1(1/2), 85–118 (1996)
16. Tarantola, A.: Inverse Problem Theory and Methods for Model Parameter Estimation. SIAM, Philadelphia, PA, USA (2004)