Approximate Bayesian Computation for Partially Identified Models

Alvarez, Luis Antonio

20 March 2023

Approximate Bayesian Computation for Partially Identified Models

Luis Antonio Fantozzi Alvarez *

March 2023

Abstract

Partial identification is a prominent feature of several economic models. Such prevalence has spurred a large literature on valid set estimation under partial identification from a frequentist viewpoint. From the Bayesian perspective, it is well known that, under partial identification, the asymptotic validity of Bayesian credible sets in conducting frequentist inference, which is ensured by several Bernstein von-Mises theorems available in the literature, breaks down. Existing solutions to this problem require either knowledge of the map between the distribution of the data and the identified set – which is generally unavailable in more complex models –, or modifications to the methodology that difficult the Bayesian interpretability of the proposed solution. In this paper, I show how one can leverage Approximate Bayesian Computation, a Bayesian methodology designed for settings where evaluation of the model likelihood is unfeasible, to reestablish the asymptotic validity of Bayesian credible sets in conducting frequentist inference, whilst preserving the core interpretation of the Bayesian approach and dispensing with knowledge of the map between data and identified set. Specifically, I show in a simple, yet encompassing, setting how, by calibrating the main tuning parameter of the ABC methodology, one could hope to achieve asymptotic frequentist coverage. Based on my findings, I then propose a semiautomatic algorithm for selecting this parameter and constructing valid confidence sets.

This is a work in progress. In future versions, I intend to present further theoretical results, Monte Carlo simulations and an empirical application on the Economics of Networks.

*Post-doctoral researcher, Fundação Getúlio Vargas. email: luisfantozzialvarez@gmail.com.

1 Introduction

Partially identified models are ubiquitous in Economics (Ho and Rosen, 2017; Molinari, 2020). In several settings, point identification is either (i) difficult to establish; or (ii) requires stringent conditions that may not be credible. In these settings, a partial identification analysis may produce more credible estimates and still be informative (Tamer, 2010, 2019). Interval estimation under partial identification has been extensively studied in the frequentist framework. I refer the reader to the surveys of Canay and Shaikh (2017) and Molinari (2020) for comprehensive expositions on the topic.

From the Bayesian viewpoint, inference under partial identification has been somewhat less studied. Gustafson (2010, 2012, 2014) has shown in different settings that, when point identification fails, the choice of prior plays a crucial role on posterior inferences, even in large samples.[1] Relatedly, Moon and Schorfheide (2012) show that, in partially identified models, the asymptotic validity of Bayesian credible intervals in conducting *frequentist* inference, which is ensured by several "Bernstein von-Mises" theorems available in the literature, breaks down. The intuition for this phenomenon is provided by the results in Gustafson (2010, 2012), which show that, under partial identification, the posterior is asymptotically supported on the identified set; and Bayesian credible sets with nontrivial posterior coverage will be eventually strictly contained in this region. In these settings, frequentist validity of credible intervals will be a feature essentially of the *prior*, and cannot be ensured to hold for every possible value in the identified set.

Since the results of Moon and Schorfheide (2012) were published, there have been several proposals in the literature which aim to reconcile Bayesian and frequentist inference in partially identified models. Müller and Norets (2016) introduce an algorithm for modifying a prior in a Bayesian analysis so the resulting confidence set attains frequentist coverage. Their proposed methodology requires evaluation of the model likelihood, which is not a trivial task in more complex models. Their proposed methodology also departs from standard Bayesian analyses, where the prior may or should embody scientific knowledge. Kline and Tamer (2016) recommend Bayesian inference be conducted in reduced-form, point-identified parameters, and that inferene on the *identified set* be made by pushing draws from the posterior of the reduced form through the map that associates, for every value of the reduced form, the corresponding identified set. The resulting posterior distribution of identified sets can then be used to conduct inference on features of the *identified set*, whilst preserving asymptotic frequentist validity. Their approach suffers from some

[1]See Wechsler et al. (2013) for an instructive example.

drawbacks, though. First, the authors' suggested construction of credible sets can be quite conservative. Moreover, their method requires knowledge of the map between the reduced form and the identified set, which is usually unknown in more complex applications. Finally, in settings where structural parameters can be identified with a scientific theory, a prior for the structural parameter appears "more principled" than one on the reduced form. Drawing on the "push-through" insight of Kline and Tamer (2016), Liao and Simoni (2019) propose computationally efficient methods to draw posterior samples from the identified set and construct credible regions in smooth convex models. Though less conservative than Kline and Tamer's approach, their method still suffers from the remaining drawbacks. Giacomini and Kitagawa (2021) propose that Bayesian inference be conducted in partially identified models in the standard way, but that sensitivity to the choice of prior be analysed by varying the prior of the structural parameter over a nonparametric class. The authors propose constructing confidence sets that achieve posterior coverage in the "worst-prior case", and show this procedure leads to asymptotically valid frequentist coverage. While their method achieves correct frequentist coverage, it departs from standard Bayesian analyses, where one sticks with a single prior and credible sets can be easily constructed using the resulting posterior draws. Their approach also requires that the map between reduced form and identified set be known.

In this paper, I propose an alternative solution that seeks to reestablish frequentist validity of Bayesian credible sets in large samples, whilst overcoming some of the issues previously discussed. Through a simple, yet encompassing example, I show that, by using Approximate Bayesian Computation (ABC), a class of algorithms originally proposed in settings where evaluation of the model likelihood is unfeasible (Beaumont et al., 2002; Tavaré, 2018), one is able to conduct standard Bayesian analysis and still retain asymptotic validity in the frequentist sense. Specifically, I show that, by carefully choosing the main tuning parameter required by the algorithm, the resulting approximate high-posterior density sets will have asymptotic valid frequentist coverage. Importantly, due to the likelihood-free nature of the method, we need not know the map between reduced form and identified set. Based on my findings, I then propose an algorithm to semiautomatically calibrate the tuning paramaters. With these tools, I hope to provide a more unified approach between Bayesian and frequentist inference in partially identified models.

To get the intuition behind the main result, it is instructive to review the ABC approach. In its simplest inception, ABC is an accept-reject algorithm that presumes the researcher be able to draw a vector of statistics $T \in \mathbb{R}^d$ from the data generating process under structural parameter θ. Specifically, given a draw θ from prior π, the researcher draws $T \sim P_\theta$, where P_θ is the distribution function under θ. The researcher then accepts θ if $\|\hat{T}-T\| < \epsilon$, where $\hat{T}$ is the vector of statistics computed under the data and $\epsilon > 0$ is a pre-specified tolerance. The resulting posterior draws can be seen as coming from:

$$p^{\epsilon}(\theta|\hat{T}) \propto P_{\theta,T\sim}[\|T-\hat{T}\| < \epsilon]\pi(\theta)\,. \tag{1}$$

As, $\epsilon \to 0$, the distribution converges to the Bayesian posterior that uses the likelihood of T given θ. More complex implementations of the ABC approach seek to provide a better-quality approximation than the simple accept-reject algorithm, but retain the essential feature of approximating the posterior by accepting draws with some tolerance (Sisson et al., 2018).

The choice of tolerance ϵ is crucial for proper functioning of the ABC approach. Unsurprisingly, it has been extensively studied in the literature (Beaumont, 2019). Blum (2010) shows that, with a finite number of simulations, there is a bias-variance tradeoff in the choice of ϵ akin to the one found in frequentist kernel density estimation. Li and Fearnhead (2018a,b) and Frazier et al. (2018) study the frequentist properties of ABC. They show that, in point identified models, choices of ϵ that are optimal for point estimation lead the posterior to overstate frequentist uncertainty. It is precisely this insight we seek to extend and explore in the partially identified setting. Since the failure of credible intervals in achieving frequentist coverage can be attributed to the posterior being asymptotically supported in the identified set, we show that, by carefully choosing ϵ, we can detain this effect, insasmuch that high-posterior density credible sets attain frequentist coverage. Importantly, our proposed method for calibrating ϵ does not require knowledge of the map between reduced form and identified set, and is agnostic about the ABC method used for posterior draws. Thus, our approach can be used in conjunction with the main computational methods in the ABC literature (Marin et al., 2011). Relatedly, even though the construction of high-posterior density sets is complicated by the fact that the likelihood is unknown, any method that estimates this density from posterior draws can be used, provided they yield good enough approximations.

The remainder of this paper is structured as follows. Section 2 illustrates the main insights on the relation between partial identification and ABC by means of a simple example. Based on these findings, Section 3 proposes an algorithm for the calibration of the tolerance ϵ that enables ABC to attain frequentist coverage. Section 4 concludes. This is a work in progress. Future versions of this working paper will feature a more general theory, Monte Carlo simulations and an empirical application on the Economics of Networks.

2 A motivating example

I start with a simple, yet encompassing, example to motivate the main results. The more general theory, where the model likelihood need not be exactly Gaussian and multidimensional summary statistics are considered, is a work in progress and deferred to future versions of this working paper. Our starting point is a measure space (Θ, Ω, μ), where Θ denotes the structural parameter space of our model. We make no assumptions on the measure space, beyond the requirement that the Bayes rule produces a valid posterior distribution in this setting (see Ghosal and Vaart, 2017, for a discussion). For each $n \in \mathbb{N}$, we associate with the structural parameter space a family of distributions on $\mathbb{R}^d$, $\mathcal{P}_n = \{P_{n,\theta} : \theta \in \Theta\}$. The index n is usually interpreted as a sample size, but more generally it could simply indicate the closeness to an approximation (Geyer, 2013). In our simple example, we set $d = 1$ and $P_{n,\theta} = N(T(\theta), \sigma^2(\theta)/n)$. Denoting by θ_0 the true parameter value, the researcher observes:

$$\hat{T} \sim P_{n,\theta_0} \,,$$

and, given a prior density π on (Θ, Ω, μ), the posterior distribution is given by:

$$H_0(A) = \frac{\int_A \frac{1}{\sigma(\theta)} \phi\left(\sqrt{n}\left(\frac{\hat{T} - T(\theta)}{\sigma(\theta)}\right)\right) \pi(\theta)\mu(d\theta)}{\int \frac{1}{\sigma(\theta)} \phi\left(\sqrt{n}\left(\frac{\hat{T} - T(\theta)}{\sigma(\theta)}\right)\right) \pi(\theta)\mu(d\theta)}, \quad A \in \Omega\,.$$

To understand the properties of the posterior distribution under partial identification, we define the set $\mathcal{C}_T(\theta_0) = \{\theta \in \Theta : T(\theta) = T(\theta_0)\}$. Observe that this set corresponds to the identification region associated with the first moment of $\hat{T}$. In the more general theory available in a work in progress version of this paper, the likelihood need not be Gaussian, and $T(\cdot)$ is, borrowing the terminology from the Indirect Inference literature (Gouriéroux et al., 2010), the *binding function* defined by $\hat{T} \overset{P_{n,\theta}}{\to} T(\theta)$, where $\overset{P_{n,\theta}}{\to}$ denotes convergence in probability under the sequence $P_{n,\theta}$.

Suppose that $\int_{\mathcal{C}_T(\theta_0)} \pi(\theta)\mu(d\theta) > 0$.[2] In this case, straightforward application of

[2] Even if this assumption does not hold for the *full* parameter space, similar arguments may still hold for the posterior induced by transformations of the structural parameter. For example, if $\Theta = \mathbb{R}^2$, $C_T(\theta_0) = \{(t, t^2) : t \in [a, b]\}$ and one adopts a Gaussian prior on $\mathbb{R}^2$, then the assumption is not satisfied. Yet, if we consider the transformation $\tilde{\theta} = h(\theta) = \theta_1$, the induced prior will assign positive mass to the identified set of $\tilde{\theta}$. In this case, our arguments may be used to show how one can construct credible sets with valid frequentist coverage for $\tilde{\theta}$. See Moon and Schorfheide (2012) for a related discussion.

the dominated convergence theorem shows that, as $n \to \infty$:

$$\sup_{A \in \Omega} |H_0(A) - H_*(A)| \overset{a.s.}{\to} 0,$$

where

$$H_*(A) = \frac{\int_{A \cap C_T(\theta_0)} \frac{1}{\sigma(\theta)} \phi\left(\frac{\sigma(\theta_0)}{\sigma(\theta)} Z\right) \pi(\theta)\mu(d\theta)}{\int_{C_T(\theta_0)} \frac{1}{\sigma(\theta)} \phi\left(\frac{\sigma(\theta_0)}{\sigma(\theta)} Z\right) \pi(\theta)\mu(d\theta)},$$

and $Z = \sqrt{n}(\hat{T} - T(\theta_0))/\sigma(\theta_0)$. This is a restatement of the results of Gustafson (2010) and Moon and Schorfheide (2012), to our setting. It suggests that, from a frequentist perspective, Bayesian credible intervals constructed under partial identification will generally undercover the parameter of interest, as the posterior is asymptotically supported on (a subset of) the identified set. Indeed, letting A_n be a sequence of credible sets with posterior coverage bounded above by strictly less than unity, we have from the convergence above that:

$$\lim_{n \to \infty} \mathbb{P}_{n,\theta_0}[A_n \subsetneq C_T(\theta_0)] = 1\,,$$

i.e. the credible interval is eventually strictly contained in the identified set, $\mathbb{P}_{n,\theta_0}$ almost surely. In addition, if we consider high-posterior density sets of the form

$$A_n = \left\{\theta \in \Theta : \frac{1}{\sigma(\theta)} \phi\left(\sqrt{n}\left(\frac{\hat{T} - T(\theta)}{\sigma(\theta)}\right)\right) \pi(\theta) > c_{1-\alpha}(\hat{T})\right\},$$

where the constant is chosen as to ensure $(1-\alpha)$-posterior coverage; and we assume $\sigma(\theta)$ to be constant in $C_T(\theta_0)$, meaning that there is no additional identifying information in the second moment of $\hat{T}$ (see footnote 3 below for a discussion), we expect from our results that:

$$\lim_{n \to \infty} \mathbb{P}_{n,\theta_0}[\theta_0 \in A_n] = \mathbf{1}\{\pi(\theta_0) > c^*\}\,,$$

where c^* is the largest threshold such that $\int_{\mathcal{C}_T(\theta_0)} \mathbf{1}\{\pi(\theta) > c^*\}\pi(\theta)\mu(d\theta) \geq (1 - \alpha) \int_{\mathcal{C}_T(\theta_0)} \pi(\theta)\mu(d\theta)$. In this setting, correct coverage is essentially a feature of the prior, and cannot be ensured to hold for every possible value of the structural parameter in the identified set, i.e.

$$\inf_{\theta \in \Theta_0} \lim_{n \to \infty} \mathbb{P}_{n,\theta}[\theta \in \mathcal{A}_n] = 0\,.$$

Given that the difficulty of correct coverage lies in the posterior being strictly contained in the prior, consider "inflating" the likelihood by considering the posterior:

$$H_\epsilon(A) = \frac{\int_A \left(\Phi\left(\sqrt{n}\left(\frac{\hat{T}-T(\theta)+\epsilon}{\sigma(\theta)}\right)\right) - \Phi\left(\sqrt{n}\left(\frac{\hat{T}-T(\theta)-\epsilon}{\sigma(\theta)}\right)\right)\right)\pi(\theta)\mu(d\theta)}{\int \left(\Phi\left(\sqrt{n}\left(\frac{\hat{T}-T(\theta)+\epsilon}{\sigma(\theta)}\right)\right) - \Phi\left(\sqrt{n}\left(\frac{\hat{T}-T(\theta)-\epsilon}{\sigma(\theta)}\right)\right)\right)\pi(\theta)\mu(d\theta)}.$$

This is precisely the posterior implemented by the ABC algorithm discussed in the introduction. Observe that, as $n \to \infty$:

$$\sup_{A\in\Omega} |H_\epsilon(A) - H^*_\epsilon(A)| \overset{a.s.}{\to} 0,$$

where

$$\begin{aligned} H^*_\epsilon(A) \propto \int_{A\cap B_{\epsilon,T}(\theta_0)} \pi(\theta)\mu(d\theta_0)+ \\ \int_{A\cap R_{\epsilon,T}(\theta_0)} \Phi\left(\frac{\sigma(\theta_0)}{\sigma(\theta)}Z\right)\pi(\theta)\mu(d\theta_0)+ \\ \int_{A\cap R_{-\epsilon,T}(\theta_0)} \left(1-\Phi\left(\frac{\sigma(\theta_0)}{\sigma(\theta)}Z\right)\right)\pi(\theta)\mu(d\theta_0)\,, \end{aligned} \tag{2}$$

with $B_{\epsilon,T}(\theta_0) = \{\theta \in \Theta : |T(\theta) - T(\theta_0)| < \epsilon\}$, and $R_{s,T}(\theta_0) = \{\theta \in \Theta : T(\theta) = T(\theta_0) + s\}$. This result shows that, with fixed ϵ, the ABC posterior is asymptotically supported on a set larger than $C_T(\theta_0)$. By carefully choosing ϵ, we hope to construct confidence sets that yield valid frequentist coverage. To see this intuitively, we consider high-posterior density sets in the asymptotic posterior. Specifically, denoting by $p^*_\epsilon(\cdot|Z)$ the density of H^*_ϵ, we consider credible sets constructed as:

$$\mathcal{S} = \{\theta \in \Theta : p^*_\epsilon(\theta|Z) > c_{1-\alpha}(Z)\}, \tag{3}$$

where the threshold $c_{1-\alpha}(Z)$ is chosen such that $H^*_\epsilon(\mathcal{S}) \geq (1-\alpha)$. Suppose that $\sigma(\theta) = \sigma(\theta_0)$ for all $\theta \in R_{-\epsilon,T(\theta_0)} \cup R_{\epsilon,T(\theta_0)}$. Notice that this assumption limits identifying information available in second moments.[3] In this case, for $\mathcal{S}$ to achieve

[3]This can be ensured by augmenting $\hat{T}$ with further statistics, so that second moments become uninformative. Alternatively, this condition may be forced to hold by working with *studentised* statistics, though in this case the identification region $C_T(\theta_0)$ will in general differ from the unstudentised setting. We defer discussion on these different strategies for future versions of this working paper.

frequentist coverage, it suffices that: first, with probability at least $(1-\alpha)$,

$$\pi(\theta) > \Phi(Z)\pi(\theta') \vee (1-\Phi(Z))\pi(\theta''), \quad \forall \theta \in C_T(\theta_0), \theta' \in R_{\epsilon,T}(\theta_0), \theta'' \in R_{-\epsilon,T}(\theta_0)\,,$$

which ensures the posterior density assigned to the identified set is among the largest with probability at least $(1-\alpha)$. Since $\Phi(Z) \sim \text{Uniform}[0,1]$, the condition is equivalent to:

$$\mathbb{P}[U\bar{\pi}^{\epsilon} \vee (1-U)\bar{\pi}^{-\epsilon} \leq \underline{\pi}_0] \geq 1-\alpha\,, \tag{4}$$

where $U := \Phi(Z)$ and $\bar{\pi}^s = \sup_{\theta \in R_{s,T}(\theta_0)} \pi(\theta)$ and $\underline{\pi}^s = \inf_{\theta \in R_{s,T}(\theta_0)} \pi(\theta)$.

The second required condition limits the relative mass of the region $B_{\epsilon,T}(\theta_0)$, on the event $U\bar{\pi}^{\epsilon} \vee (1-U)\bar{\pi}^{-\epsilon} \leq \underline{\pi}_0$. This ensures that, on this event, the identified set will be always contained within the credible set. Specifically, we require that, on $U\bar{\pi}^{\epsilon} \vee (1-U)\bar{\pi}^{-\epsilon} \leq \underline{\pi}_0$:

$$\begin{aligned} \int_{B_{\epsilon,T}(\theta_0)} \pi(\theta)\mu(d\theta_0) \leq (1-\alpha)\Big[\int_{B_{\epsilon,T}(\theta_0)} \pi(\theta)\mu(d\theta_0) + \\ \int_{R_{\epsilon,T}(\theta_0)} \Phi\left(Z\right)\pi(\theta)\mu(d\theta_0) + \\ \int_{R_{-\epsilon,T}(\theta_0)} \left(1-\Phi\left(Z\right)\right)\pi(\theta)\mu(d\theta_0)\Big] \text{ a.s.} \end{aligned}$$

This condition can be seen to be equivalent to:

$$\begin{aligned} \frac{\alpha}{(1-\alpha)} \int_{B_{\epsilon,T}(\theta_0)} \pi(\theta)\mu(d\theta_0) \leq \\ \Big[\int_{R_{\epsilon,T}(\theta_0)} \Big(1 - \frac{\underline{\pi}_0}{\bar{\pi}^{-\epsilon}}\Big) \pi(\theta)\mu(d\theta_0) + \int_{R_{-\epsilon,T}(\theta_0)} \frac{\underline{\pi}_0}{\bar{\pi}^{-\epsilon}} \pi(\theta)\mu(d\theta_0)\Big] \wedge \\ \Big[\int_{R_{\epsilon,T}(\theta_0)} \frac{\underline{\pi}_0}{\bar{\pi}^{\epsilon}} \pi(\theta)\mu(d\theta_0) + \int_{R_{-\epsilon,T}(\theta_0)} (1 - \frac{\underline{\pi}_0}{\bar{\pi}^{\epsilon}}) \pi(\theta)\mu(d\theta_0)\Big]\,. \end{aligned} \tag{5}$$

In practice, these conditions cannot be immediately used to calibrate ϵ, because the regions $C_T(\theta_0)$ and $R_c(\theta_0)$ are not known. However, as our proposed algorithm in Section 3 shows, we can conservatively estimate the unknown quantities so asymptotic frequentist coverage is achieved.

3 Proposed algorithm

In light of the preceding discussion, Algorithm 1 introduces a semiautomatic procedure for the calibration of ϵ. The algorithm searches for a candidate tolerance $\epsilon \in [\underline{\epsilon}, \overline{\epsilon}]$, by verifying whether estimated versions of (4) and (5) hold. The interval $[\underline{\epsilon}, \overline{\epsilon}]$ may be chosen using the standard logic in the ABC literature (e.g. Li and Fearnhead, 2018b), where $\underline{\epsilon}$ and $\overline{\epsilon}$ correspond, respectively, to the smallest and largest tolerable acceptance rates, given the number of Monte Carlo draws S. The algorithm also depends on a bandwidth $b > 0$ for the estimation of the regions $R_{T,c}(\theta_0)$. In future versions, I intend to explore the correct tuning of this hyperparameter, as well as to provide formal continuity conditions on π and that ensure validity of Algorithm 1.

Once a tolerance $\epsilon^* > 0$ is found, the researcher may use any ABC algorithm that targets the approximate posterior (1).[4] Using the approximate posterior draws, the researcher may then choose a semi- or nonparametric method to estimate the posterior density; and use this estimated density to construct the high posterior density sets.

[4]It is immediate to extend the proposed algorithm to settings where acceptance of a draw occurs with probability $K(\|\hat{T} - \tilde{T}\|/\epsilon)$, with K being a rescaled kernel. This is a common approach in the ABC literature (Fearnhead and Prangle, 2012; Li and Fearnhead, 2018b).

Algorithm 1 Proposed algorithm for the calibration of the tolerance in ABC

Require: $S \in \mathbb{N}$, $\underline{\epsilon} > 0$, $\overline{\epsilon} > 0$, $b > 0$, step > 0.

Ensure: $\epsilon^* \in [\underline{\epsilon}, \overline{\epsilon}]$

$d \leftarrow []$

for s=1 to S **do**

Draw $\theta_s \sim \pi$.

Draw $\tilde{T}_s \sim P_{n,\theta_s}$.

$d \leftarrow [d, (\theta_s, \tilde{T}_s)]$.

end for

Estimate $\hat{\underline{\pi}}_0 = \min_{s:|\tilde{T}_s - \hat{T}| \leq \underline{\epsilon}} \pi(\theta_s)$.

$\epsilon^* \leftarrow \overline{\epsilon}$.

while $\epsilon^* \geq \underline{\epsilon}$ and stop criterion not met **do**

Estimate $\widehat{\int_{B_{\epsilon^*,T}(\theta_0)} \pi(\theta)}\mu(d\theta_0) = \frac{1}{S}\sum_{s=1}^{S} \mathbf{1}\{|\hat{T} - \tilde{T}_s| \leq \epsilon^*\}$.

Estimate $\widehat{\int_{R_{\epsilon^*,T}(\theta_0)} \pi(\theta)}\mu(d\theta_0) = \frac{1}{S}\sum_{s=1}^{S} \mathbf{1}\{\hat{T} + \epsilon^* - b \leq \tilde{T}_s \leq \hat{T} + \epsilon^* + b\}$ and $\widehat{\int_{R_{-\epsilon^*,T}(\theta_0)} \pi(\theta)}\mu(d\theta_0) = \frac{1}{S}\sum_{s=1}^{S} \mathbf{1}\{\hat{T} - \epsilon^* - b \leq \tilde{T}_s \leq \hat{T} - \epsilon^* + b\}$.

Estimate $\widehat{\overline{\pi}^{\epsilon^*}} = \max_{s:\hat{T}+\epsilon^*-b \leq \tilde{T}_s \leq \hat{T}+\epsilon^*+b} \pi(\theta_s)$ and $\widehat{\overline{\pi}^{-\epsilon^*}} = \max_{s:\hat{T}-\epsilon^*-b \leq \tilde{T}_s \leq \hat{T}-\epsilon^*+b} \pi(\theta_s)$.

if ϵ^* satisfies (4) and (5) **then**

stop

else

$\epsilon^* \leftarrow \epsilon^* -$ step.

end if

end while

4 Concluding remarks

This paper showed, by the means of a simple example, that Approximate Bayesian Computation, a popular algorithm in settings where the model likelihood is intractable, can be a convenient method to conduct Bayesian inference in partially identified models. Specifically, I show that, by properly calibrating the main tuning parameter in the algorithm, one can construct credible sets with asymptotic frequentist validity.

This is a work in progress. In future versions, I intend to incorporate the general theory, which does not assume a normal likelihood nor unidimensional sufficient statistics, and to provide high-level conditions on the prior density and on the estimator of the posterior density for validity of the proposed method. I also intend to present Monte Carlo simulations, and an empirical application on the Economics of Networks.

References

Beaumont, Mark A (2019). "Approximate Bayesian Computation". In: *Annual Review of Statistics and Its Application* 6.1, pp. 379–403. DOI: 10.1146/annurev-statistics-030718-105212. eprint: https://doi.org/10.1146/annurev-statistics-030718-105212. URL: https://doi.org/10.1146/annurev-statistics-030718-105212.

Beaumont, Mark A, Wenyang Zhang, and David J Balding (Dec. 2002). "Approximate Bayesian Computation in Population Genetics". In: *Genetics* 162.4, pp. 2025–2035. ISSN: 1943-2631. DOI: 10.1093/genetics/162.4.2025. eprint: https://academic.oup.com/genetics/article-pdf/162/4/2025/42049447/genetics2025.pdf. URL: https://doi.org/10.1093/genetics/162.4.2025.

Blum, Michael G. B. (2010). "Approximate Bayesian Computation: A Nonparametric Perspective". In: *Journal of the American Statistical Association* 105.491, pp. 1178–1187. DOI: 10.1198/jasa.2010.tm09448. eprint: https://doi.org/10.1198/jasa.2010.tm09448. URL: https://doi.org/10.1198/jasa.2010.tm09448.

Canay, Ivan A. and Azeem M. Shaikh (2017). "Practical and Theoretical Advances in Inference for Partially Identified Models". In: *Advances in Economics and Econometrics: Eleventh World Congress*. Ed. by Bo Honoré et al. Vol. 2. Econometric Society Monographs. Cambridge University Press, pp. 271–306. DOI: 10.1017/9781108227223.009.

Fearnhead, Paul and Dennis Prangle (2012). "Constructing summary statistics for approximate Bayesian computation: semi-automatic approximate Bayesian computation". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 74.3, pp. 419–474. DOI: https://doi.org/10.1111/j.1467-9868.2011.01010.x. eprint: https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-9868.2011.01010.x. URL: https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9868.2011.01010.x.

Frazier, D T et al. (June 2018). "Asymptotic properties of approximate Bayesian computation". In: *Biometrika* 105.3, pp. 593–607. ISSN: 0006-3444. DOI: 10.1093/biomet/asy027. eprint: https://academic.oup.com/biomet/article-pdf/105/3/593/25470059/asy027.pdf. URL: https://doi.org/10.1093/biomet/asy027.

Geyer, Charles J. (2013). "Asymptotics of Maximum Likelihood without the LLN or CLT or Sample Size Going to Infinity". In: *Advances in Modern Statistical Theory and Applications: A Festschrift in honor of Morris L. Eaton*. Institute of Mathematical Statistics, pp. 1–24. DOI: 10.1214/12-imscoll1001. URL: https://doi.org/10.1214/12-imscoll1001.

Ghosal, Subhashis and Aad van der Vaart (2017). "Introduction". In: *Fundamentals of Nonparametric Bayesian Inference*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, pp. 1–9. DOI: 10.1017/9781139029834.002.

Giacomini, Raffaella and Toru Kitagawa (2021). "Robust Bayesian Inference for Set-Identified Models". In: *Econometrica* 89.4, pp. 1519–1556. DOI: https://doi.org/10.3982/ECTA16773. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA16773. URL: https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA16773.

Gouriéroux, Christian, Peter C.B. Phillips, and Jun Yu (2010). "Indirect inference for dynamic panel models". In: *Journal of Econometrics* 157.1. Nonlinear and Nonparametric Methods in Econometrics, pp. 68–77. ISSN: 0304-4076. DOI: https://doi.org/10.1016/j.jeconom.2009.10.024. URL: https://www.sciencedirect.com/science/article/pii/S0304407609002772.

Gustafson, Paul (2010). "Bayesian Inference for Partially Identified Models". In: *The International Journal of Biostatistics* 6.2. DOI: doi:10.2202/1557-4679.1206. URL: https://doi.org/10.2202/1557-4679.1206.

— (2012). "On the behaviour of Bayesian credible intervals in partially identified models". In: *Electronic Journal of Statistics* 6.none, pp. 2107–2124. DOI: 10.1214/12-EJS741. URL: https://doi.org/10.1214/12-EJS741.

— (2014). "Bayesian inference in partially identified models: Is the shape of the posterior distribution useful?" In: *Electronic Journal of Statistics* 8.1, pp. 476–496. DOI: 10.1214/14-EJS891. URL: https://doi.org/10.1214/14-EJS891.

Ho, Kate and Adam M. Rosen (2017). "Partial Identification in Applied Research: Benefits and Challenges". In: *Advances in Economics and Econometrics: Eleventh World Congress*. Ed. by Bo Honoré et al. Vol. 2. Econometric Society Monographs. Cambridge University Press, pp. 307–359. DOI: 10.1017/9781108227223.010.

Kline, Brendan and Elie Tamer (2016). "Bayesian inference in a class of partially identified models". In: *Quantitative Economics* 7.2, pp. 329–366. DOI: https://doi.org/10.3982/QE399. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/QE399. URL: https://onlinelibrary.wiley.com/doi/abs/10.3982/QE399.

Li, Wentao and Paul Fearnhead (Jan. 2018a). "Convergence of regression-adjusted approximate Bayesian computation". In: *Biometrika* 105.2, pp. 301–318. ISSN: 0006-3444. DOI: 10.1093/biomet/asx081. eprint: https://academic.oup.com/biomet/article-pdf/105/2/301/24845524/asx081.pdf. URL: https://doi.org/10.1093/biomet/asx081.

— (Jan. 2018b). "On the asymptotic efficiency of approximate Bayesian computation estimators". In: *Biometrika* 105.2, pp. 285–299. ISSN: 0006-3444. DOI: 10.1093/biomet/asx078. eprint: https://academic.oup.com/biomet/article-pdf/105/2/285/24820949/asx078.pdf. URL: https://doi.org/10.1093/biomet/asx078.

Liao, Yuan and Anna Simoni (2019). "Bayesian inference for partially identified smooth convex models". In: *Journal of Econometrics* 211.2, pp. 338–360. ISSN: 0304-4076. DOI: https://doi.org/10.1016/j.jeconom.2019.03.001. URL: https://www.sciencedirect.com/science/article/pii/S0304407619300399.

Marin, Jean-Michel et al. (Oct. 2011). "Approximate Bayesian computational methods". In: *Statistics and Computing* 22.6, pp. 1167–1180. DOI: 10.1007/s11222-011-9288-2. URL: https://doi.org/10.1007/s11222-011-9288-2.

Molinari, Francesca (2020). "Chapter 5 - Microeconometrics with partial identification". In: *Handbook of Econometrics, Volume 7A*. Ed. by Steven N. Durlauf et al. Vol. 7. Handbook of Econometrics. Elsevier, pp. 355–486. DOI: https://doi.org/10.1016/bs.hoe.2020.05.002. URL: https://www.sciencedirect.com/science/article/pii/S1573441220300027.

Moon, Hyungsik Roger and Frank Schorfheide (2012). "Bayesian and Frequentist Inference in Partially Identified Models". In: *Econometrica* 80.2, pp. 755–782. DOI: https://doi.org/10.3982/ECTA8360. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA8360. URL: https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA8360.

Müller, Ulrich K. and Andriy Norets (2016). "Coverage Inducing Priors in Nonstandard Inference Problems". In: *Journal of the American Statistical Association* 111.515, pp. 1233–1241. DOI: 10.1080/01621459.2015.1086654. eprint: https://doi.org/10.1080/01621459.2015.1086654. URL: https://doi.org/10.1080/01621459.2015.1086654.

Sisson, Scott A, Yanan Fan, and Mark A Beaumont (2018). "Overview of ABC". In: *Handbook of approximate Bayesian computation*. Chapman and Hall/CRC, pp. 3–54.

Tamer, Elie (2010). "Partial Identification in Econometrics". In: *Annual Review of Economics* 2.1, pp. 167–195. DOI: 10.1146/annurev.economics.050708.143401. eprint: https://doi.org/10.1146/annurev.economics.050708.

143401. URL: https://doi.org/10.1146/annurev.economics.050708.143401.
Tamer, Elie (2019). "The ET Interview: Professor Charles Manski". In: *Econometric Theory* 35.2, pp. 233–294. DOI: 10.1017/S0266466618000075.
Tavaré, Simon (2018). "On the history of ABC". In: *Handbook of Approximate Bayesian Computation*. Chapman and Hall/CRC, pp. 55–69.
Wechsler, Sergio, Rafael Izbicki, and Luís Gustavo Esteves (2013). "A Bayesian Look at Nonidentifiability: A Simple Example". In: *The American Statistician* 67.2, pp. 90–93. ISSN: 00031305. URL: http://www.jstor.org/stable/24591438 (visited on 02/08/2023).