# On the gap between ess(f) and cnf_size(f) 

Lisa Hellerstein, Devorah Kletenik*<br>Polytechnic Institute of NYU, 6 Metrotech Center, Brooklyn, NY, 11201, United States

## ARTICLE INFO

Article history:
Received 12 June 2011
Accepted 8 July 2012
Available online 24 August 2012

## Keywords:

DNF
CNF
$\operatorname{ess}(f)$
Horn functions
Formula size


#### Abstract

Given a Boolean function $f$, the quantity ess $(f)$ denotes the largest set of assignments that falsify $f$, no two of which falsify a common implicate of $f$. Although ess $(f)$ is clearly a lower bound on cnf_size(f) (the minimum number of clauses in a CNF formula for $f$ ), Čepek et al. showed it is not, in general, a tight lower bound [6]. They gave examples of functions $f$ for which there is a small gap between ess $(f)$ and $c n f$ _size( $f$ ). We demonstrate significantly larger gaps. We show that the gap can be exponential in $n$ for arbitrary Boolean functions, and $\Theta(\sqrt{n})$ for Horn functions, where $n$ is the number of variables of $f$. We also introduce a natural extension of the quantity $\operatorname{ess}(f)$, which we call $\operatorname{ess}_{k}(f)$, which is the largest set of assignments, no $k$ of which falsify a common implicate of $f$.


© 2012 Elsevier B.V. All rights reserved.

## 1. Introduction

Determining the smallest CNF formula for a given Boolean function $f$ is a difficult problem that has been studied for many years. (See [7] for an overview of relevant literature.) Recently, Čepek et al. introduced a combinatorial quantity, ess( $f$ ), which lower bounds cnf_size( $f$ ), the minimum number of clauses in a CNF formula representing $f$ [6]. The quantity ess( $f$ ) is equal to the size of the largest set of falsepoints of $f$, no two of which falsify the same implicate of $f$. ${ }^{1}$

For certain subclasses of Boolean functions, such as the monotone (i.e., positive) functions, $\operatorname{ess}(f)$ is equal to cnf_size( $f$ ). However, Čepek et al. demonstrated that there can be a gap between ess $(f)$ and $c n f$ _size $(f)$. They constructed a Boolean function $f$ on $n$ variables such that there is a multiplicative gap of size $\Theta(\log n)$ between cnf_size $(f)$ and ess $(f) .{ }^{2}$ Their constructed function $f$ is a Horn function. Their results leave open the possibility that ess $(f)$ could be a close approximation to cnf_size(f).

We show that this is not the case. We construct a Boolean function $f$ on $n$ variables such that there is a multiplicative gap of size $2^{\Theta(n)}$ between $c n f$ _size $(f)$ and $\operatorname{ess}(f)$. Note that such a gap could not be larger than $2^{n-1}$, since cnf_size $(f) \leq 2^{n-1}$ for all functions $f$ on $n>1$ variables.

We also construct a Horn function $f$ such that there is a multiplicative gap of size $\Theta(\sqrt{n})$ between cnf_size $(f)$ and ess $(f)$. We show that no gap larger than $\Theta(n)$ is possible.

If one expresses the gaps as a function of $c n f$ _size $(f)$, rather than as a function of the number of variables $n$, then the gap we obtain with both the constructed non-Horn and Horn functions $f$ is $c n f$ _size $(f)^{1 / 3}$. Clearly, no gap larger than cnf_size ( $f$ ) is possible.

We briefly explore a natural generalization of the quantity $\operatorname{ess}(f)$, which we call $\operatorname{ess}_{k}(f)$, which is the largest set of falsepoints, no $k$ of which falsify a common implicate of $f$. The quantity ess $(f) /(k-1)$ is a lower bound on cnf_size(f), for any $k \geq 2$.

The above results concern the size of CNF formulas. Analogous results hold for DNF formulas by duality.

[^0]
## 2. Preliminaries

### 2.1. Definitions

A Boolean function $f\left(x_{1}, \ldots, x_{n}\right)$ is a mapping $\{0,1\}^{n} \rightarrow\{0,1\}$. (Where it does not cause confusion, we often use the word "function" to refer to a Boolean function.) A variable $x_{i}$ and its negation $\neg x_{i}$ are literals (positive and negative respectively). A clause is a disjunction $(\vee)$ of literals. A term is a conjunction $(\wedge)$ of literals. A CNF (conjunctive normal form) formula is a formula of the form $c_{0} \wedge c_{1} \wedge \cdots c_{k}$, where each $c_{i}$ is a clause. A $D N F$ (disjunctive normal form) formula is a formula of the form $t_{0} \vee t_{1} \vee \cdots t_{k}$, where each $t_{i}$ is a term.

A clause $c$ containing variables from $X_{n}=\left\{x_{1}, \ldots, x_{n}\right\}$ is an implicate of $f$ if for all $x \in\{0,1\}^{n}$, if $c$ is falsified by $x$ then $f(x)=0$. A term $t$ containing variables from $X_{n}$ is an implicant of function $f\left(x_{1}, \ldots, x_{n}\right)$ if for all $x \in\{0,1\}^{n}$, if $t$ is satisfied by $x$ then $f(x)=1$.

We define the size of a CNF formula to be the number of its clauses, and the size of a DNF formula to be the number of its terms.

Given a Boolean function $f$, cnf_size $(f)$ is the size of the smallest CNF formula representing $f$. Analogously, dnf_size ( $f$ ) is the size of the smallest DNF formula representing $f$. If $f$ is the identically false function, the CNF representation of $f$ is be the empty clause and the DNF representation is $x_{1} \neg x_{1}$. Representations for the identically true function follow by duality. In both cases, $c n f$ _size $(f)=d n f$ _size $(f)=1$.

An assignment $x \in\{0,1\}^{n}$ is a falsepoint of $f$ if $f(x)=0$, and is a truepoint of $f$ if $f(x)=1$. We say that a clause $c$ covers a falsepoint $x$ of $f$ if $x$ falsifies $c$. A term $t$ covers a truepoint $x$ of $f$ if $x$ satisfies $t$.

A CNF formula $\phi$ representing a function $f$ forms a cover of the falsepoints of $f$, in that each falsepoint of $f$ must be covered by at least one clause of $\phi$. Further, if $x$ is a truepoint of $f$, then no clause of $\phi$ covers $x$. Similarly, a DNF formula $\phi$ representing a function $f$ forms a cover of the truepoints of $f$, in that each truepoint of $f$ must be covered by at least one term of $\phi$. Further, if $x$ is a falsepoint of $f$, then no term of $\phi$ covers $x$.

Given two assignments $x, y \in\{0,1\}^{n}$, we write $x \leq y$ if $\forall i, x_{i} \leq y_{i}$. An assignment $r$ separates two assignments $p$ and $q$ if $\forall i, p_{i}=r_{i}$ or $q_{i}=r_{i}$.

A partial function $f$ maps $\{0,1\}^{n}$ to $\{0,1, *\}$, where $*$ indicates that the value of $f$ is not defined on the assignment. A Boolean formula $\phi$ is consistent with a partial function $f$ if $\phi(a)=f(a)$ for all $a \in\{0,1\}^{n}$ where $f(a) \neq *$. If $f$ is a partial Boolean function, then cnf_size( $f$ ) and dnf_size( $f$ ) are the size of the smallest CNF and DNF formulas consistent with the $f$, respectively.

A Boolean function $f\left(x_{1}, \ldots, x_{n}\right)$ is monotone if for all $x, y \in\{0,1\}^{n}$, if $x \leq y$ then $f(x) \leq f(y)$. A Boolean function is anti-monotone if for all $x, y \in\{0,1\}^{n}$, if $x \geq y$ then $f(x) \leq f(y)$.

A DNF or CNF formula is monotone if it contains no negations; it is anti-monotone if all variables in it are negated. A CNF formula is a Horn-CNF if each clause contains at most one variable without a negation. If each clause contains exactly one variable without a negation it is a pure Horn-CNF. A Horn function is a Boolean function that can be represented by a Horn-CNF. It is a pure Horn function if it can be represented by a pure Horn-CNF. Horn functions are a generalization of anti-monotone functions, and have applications in artificial intelligence [11].

We say that two falsepoints, $x$ and $y$, of a function $f$ are independent if no implicate of $f$ covers both $x$ and $y$. Similarly, we say that two truepoints $x$ and $y$ of a function $f$ are independent if no implicant of $f$ covers both $x$ and $y$. We say that a set $S$ of falsepoints (truepoints) of $f$ is independent if all pairs of falsepoints (truepoints) in $S$ are independent.

The set covering problem is as follows: Given a ground set $A=\left\{e_{1}, \ldots, e_{m}\right\}$ of elements, a set $\&=\left\{S_{1}, \ldots, S_{n}\right\}$ of subsets of $A$, and a positive integer $k$, does there exist $s^{\prime} \subseteq \&$ such that $\bigcup_{S_{i} \in \delta^{\prime}}=s$ and $\left|\delta^{\prime}\right| \leq k$ ? Each set $S_{i} \in s$ is said to cover the elements it contains. Thus the set covering problem asks whether $A$ has a "cover" of size at most $k$.

A set covering instance is $r$-uniform, for some $r>0$, if all subsets $S_{i} \in \delta$ have size $r$.
Given an instance of the set covering problem, we say that a subset $A^{\prime}$ of ground set $A$ is independent if no two elements of $A^{\prime}$ are contained in a common subset $S_{i}$ of $ร$.

## 3. The quantity ess(f)

We begin by restating the definition of $\operatorname{ess}(f)$ in terms of independent falsepoints. We also introduce an analogous quantity for truepoints. (The notation ess ${ }^{d}$ refers to the fact that this is a dual definition.)

Definition 1. Let $f$ be a Boolean function. The quantity $\operatorname{ess}(f)$ denotes the size of the largest independent set of falsepoints of $f$. The quantity $\operatorname{ess}^{d}(f)$ denotes the largest independent set of truepoints of $f$.

As was stated above, Čepek et al. introduced the quantity ess $(f)$ as a lower bound on $c n f$ _size $(f)$. The fact that ess $(f) \leq$ cnf_size( $f$ ) follows easily from the above definitions, and from the following facts: (1) if $\phi$ is a CNF formula representing $\bar{f}$, then every falsepoint of $f$ must be covered by some clause of $\phi$, and (2) each clause of $\phi$ must be an implicate of $f$.

Let $f^{\prime}$ denote the function that is the complement of $f$, i.e. $f^{\prime}(a)=\neg f(a)$ for all assignments $a$. Since, by duality, $\operatorname{ess}\left(f^{\prime}\right)=\operatorname{ess}^{d}(f)$ and $\operatorname{cnf}$ _size $\left(f^{\prime}\right)=d n f$ _size $(f)$, it follows that ess $\left(f^{\prime}\right) \leq d n f \_$size $(f)$.

Property 1 ([6]). Two falsepoints of $f, x$ and $y$, are independent iff there exists a truepoint $a$ of $f$ that separates $x$ and $y$.
Consider the following decision problem, which we will call ESS: "Given a CNF formula representing a Boolean function $f$, and a number $k$, is $\operatorname{ess}(f) \leq k ? "$. Using Property 1 , this problem is easily shown to be in co-NP [6].

We can combine the fact that ESS is in co-NP with results on the hardness of approximating CNF-minimization, to get the following preliminary result, based on a complexity-theoretic assumption.

Proposition 1. If co-NP $\neq \Sigma_{2}^{P}$, then for some $\gamma>0$, there exists an infinite set of Boolean functions $f$ such that ess(f) $n^{\gamma}<$ cnf_size( $f$ ), where $n$ is the number of variables of $f$.

Proof. Consider the Min-CNF problem (decision version): Given a CNF formula representing a Boolean function $f$, and a number $k$, is cnf_size $(f) \leq k$ ? Umans proved that it is $\Sigma_{2}^{P}$-complete to approximate this problem to within a factor of $n^{\gamma}$, for some $\gamma>0$, where $n$ is the number of variables of $f$ [12]. (Approximating this problem to within some factor $q$ means answering "yes" whenever cnf_size $(f) \leq k$, and answering "no" whenever cnf_size $(f)>k q$. If $k<c n f$ _size $(f) \leq k q$, either answer is acceptable.)

Suppose ess $(f) n^{\gamma} \geq$ cnf_size( $f$ ) for all Boolean functions $f$. Then one can approximate Min-CNF to within a factor of $n^{\gamma}$ in co-NP by simply using the co-NP algorithm for ESS to determine whether ess $(f) \leq k$. Even if ess $(f) n^{\gamma} \geq c n f \_\operatorname{size}(f)$ for a finite set $S$ of functions, one can still approximate Min-CNF to within a factor of $n^{\gamma}$ in co-NP, by simply handling the finite number of functions in $S$ explicitly as special cases. Since approximating Min-CNF to within this factor is $\Sigma_{2}^{P}$-complete, $\Sigma_{2}^{P} \subseteq$ co-NP. By definition, co-NP $\subseteq \Sigma_{2}^{P}$, so $\Sigma_{2}^{P}=$ co-NP.

The non-approximability result of Umans for Min-CNF, used in the above proof, is expressed in terms of the number of variables $n$ of the function. Umans also showed [13] that it is $\Sigma_{2}^{P}$ complete to approximate Min-CNF to within a factor of $m^{\gamma}$, for some $\gamma \geq 0$, where $m=c n f \_$size( $f$ ). Thus we can also prove that, if NP $\neq \Sigma_{2}^{P}$, then for some $\gamma>0$, there is an infinite set of functions $f$ such that $\operatorname{ess}(f)<c n f \_\operatorname{size}(f)^{1-\gamma}$.

The assumption that $\Sigma_{2}^{P} \neq$ co-NP is not unreasonable, so we have grounds to believe that there is an infinite set of functions for which the gap between ess $(f)$ and $c n f$ _size $(f)$ is greater than $n^{\gamma}$ (or cnf_size $(f)^{\gamma}$ ) for some $\gamma$. Below, we will explicitly construct such sets with larger gaps than that of Proposition 1, and with no complexity theoretic assumptions.

We can also prove a proposition similar to Proposition 1 for Horn functions, using a different complexity theoretic assumption. (Since the statement of the proposition includes a complexity class parameterized by the standard input-size parameter $n$, we use $N$ instead of $n$ to denote the number of inputs to a Boolean function.)

Proposition 2. If NP $\nsubseteq$ co-NTIME( $n^{\text {polylog }(n)}$ ), then for some $\epsilon$ such that $0<\epsilon<1$, there exists an infinite set of Horn functions $f$ such that $\frac{\text { cnf_size(f) }}{\text { ess }(f)} \geq 2^{\log ^{1-\epsilon} N}$, where $N$ is the number of input variables of $f$.

Proof. Consider the following Min-Horn-CNF problem (decision version): Given a Horn-CNF $\phi$ representing a Horn function $f$, and an integer $k \geq 0$, is cnf_size $(f) \leq k$ ? Bhattacharya et al. [5] showed that there exists a deterministic, many-one reduction (i.e. a Karp reduction), running in time $O\left(n^{\text {polylog(n) }}\right.$ ) (where $n$ is the size of the input), from an NP-complete problem to the problem of approximating Min-Horn-CNF to within a factor of $2^{\log ^{1-\epsilon} N}$, where $N$ is the number of input variables of $f$.

Suppose that $\frac{\text { cnf_size(f) }}{\text { ess }(f)}$ is at most $2^{\log ^{1-\epsilon} N}$ for all Boolean functions $f$. It is well known that given a Horn-CNF $f$, the size of the smallest (functionally) equivalent Horn-CNF is precisely cnf_size(f). Thus given a Horn-CNF $\phi$ on $N$ variables, and a number $k$, if there does not exist a Horn-CNF equivalent to $\phi$ of size less than $2^{\log ^{1-\epsilon} N} \times k$, this can be verified non-deterministically in polynomial time (by verifying that $\operatorname{ess}(f) \geq k$ ). Thus the complement of Min-Horn-CNF is approximable to within a factor of $2^{\log ^{1-\epsilon} N}$, in deterministic time $n^{\text {polylog (n) }}$ (where $n$ is the size in bits of the input Horn-CNF, and $N$ is the number of variables in the input Horn-CNF). Combining this fact with the reduction of Bhattacharya et al. implies that the complement of an NP-complete problem can be solved in non-deterministic time $n^{\text {polylog }(n)}$. Thus NP is contained in co-NTIME $\left(n^{\text {polylog }(n)}\right)$. The same holds if $\frac{c n f \_ \text {size }(f)}{\text { ess(f) }}$ is at most $2^{\log ^{1-\epsilon} n}$ for all but a finite set of Boolean functions $f$.

## 4. Constructions of functions with large gaps between ess(f) and cnf_size(f)

We will begin by constructing a function $f$, such that $\frac{\operatorname{cnf} f \text { size }(f)}{\operatorname{ess}(f)}=\Theta(n)$. This is already a larger gap than the multiplicative gap of $\log (n)$ achieved by the construction of Čepek et al. [6], and the gap of $n^{\gamma}$ in Proposition 1 . We describe the construction of $f$, prove bounds on $c n f \_\operatorname{size}(f)$ and $\operatorname{ess}(f)$, and then prove that the ratio $\frac{c n f \_\operatorname{size}(f)}{\text { ess }(f)}=\Theta(n)$.

We will then show how to modify this construction to give a function $f$ such that $\frac{\operatorname{cnf} \text { _size }(f)}{\operatorname{ess}(f)}=2^{\Theta(n)}$, thus increasing the gap to be exponential in $n$.

At the end of this section, we will explore $\operatorname{ess}_{k}(f)$, our generalization of $\operatorname{ess}(f)$.

### 4.1. Constructing a function with a linear gap

Theorem 1. There exists a function $f\left(x_{1}, \ldots, x_{n}\right)$ such that $\frac{c n f \_ \text {size }(f)}{\operatorname{ess}(f)}=\Theta(n)$.
Proof. We construct a function $f$ such that $\frac{d n f \_ \text {size }(f)}{\text { ess }{ }^{d}(f)}=\Theta(n)$. Theorem 1 then follows immediately by duality.
Our construction relies heavily on a reduction of Gimpel from the 1960's [10], which reduces a generic instance of the set covering problem to a DNF-minimization problem. (See Czort [9] or Allender et al. [1] for more recent discussions of this reduction.)

Gimpel's reduction is as follows. Let $A=\left\{e_{1}, \ldots, e_{m}\right\}$ be the ground set of the set covering instance, and let $s$ be the set of subsets $A$ from which the cover must be formed. With each element $e_{i}$ in $A$, associate a Boolean input variable $x_{i}$. For each $S \in \ell$, let $x_{S}$ denote the assignment in $\{0,1\}^{m}$ where $x_{i}=0$ iff $e_{i} \in S$. Define the partial function $f\left(x_{1}, \ldots, x_{m}\right)$ as follows:

$$
f(x)= \begin{cases}1 & \text { if } x \text { contains exactly } m-1 \text { ones } \\ * & \text { if } x \geq x_{S} \text { and } x \text { does not contain exactly } m-1 \text { ones } \\ 0 & \text { otherwise }\end{cases}
$$

There is a DNF formula of size at most $k$ that is consistent with this partial function if and only if the elements $e_{i}$ of the set covering instance $A$ can be covered using at most $k$ subsets in $\delta$ (cf. [9]).

We apply this reduction to the simple, 2-uniform, set covering instance over $m$ elements where $\&$ consists of all subsets containing exactly two of those $m$ elements. The smallest set cover for this instance is clearly $\lceil m / 2\rceil$. The largest independent set of elements is only of size 1 , since every pair of elements is contained in a common subset of $s$. Note that this gives a ratio of minimal set cover to largest independent set of $\Theta(m)$.

Applying Gimpel's reduction to this simple set covering instance, we get the following partial function $\hat{f}$ :

$$
\hat{f}(x)= \begin{cases}1 & \text { if } x \text { contains exactly } m-1 \text { ones } \\ * & \text { if } x \text { contains exactly } m-2 \text { ones } \\ * & \text { if } x \text { contains exactly } m \text { ones } \\ 0 & \text { otherwise }\end{cases}
$$

Since the smallest set cover for the instance has size $\lceil m / 2\rceil$,

$$
d n f \_\operatorname{size}(\hat{f})=\lceil m / 2\rceil
$$

Allender et al. [1] extended the reduction of Gimpel by converting the partial function $f$ to a total function $g$. The conversion is as follows:

Let $t=m+1$ and let $s$ be the number of $*$ 's in $f(x)$. Let $y_{1}$ and $y_{2}$ be two additional Boolean variables, and let $z=z_{1} \ldots z_{t}$ be a vector of $t$ more Boolean variables. Let $S \subseteq\{0,1\}^{t}$ be a collection of $s$ vectors, each containing an odd number of 1 's (since $s \leq 2^{m}$, such a collection exists). Let $\chi$ be the function such that $\chi(x)=0$ if the parity of $x$ is even and $\chi(x)=1$ otherwise.

The total function $g$ is defined as follows:

$$
g\left(x, y_{1}, y_{2}, z\right)= \begin{cases}1 & \text { if } f(x)=1 \text { and } y_{1}=y_{2}=1 \text { and } z \in S \\ 1 & \text { if } f(x)=* \text { and } y_{1}=y_{2}=1 \\ 1 & \text { if } f(x)=*, y_{1}=\chi(x), \text { and } y_{2}=\neg \chi(x) \\ 0 & \text { otherwise } .\end{cases}
$$

Allender et al. proved that this total function $g$ obeys the following property:

$$
d n f \_\operatorname{size}(g)=s\left(d n f \_\operatorname{size}(f)+1\right)
$$

Let $\hat{g}$ be the total function obtained by setting $f=\hat{f}$ in the above definition of $g$.
We can now compute dnf_size $(\hat{g})$. Let $n$ be the number of input variables of $\hat{f}$. The total function $\hat{g}$ is defined on $n=2 m+3$ variables. Since $\operatorname{dnf} \_\operatorname{size}(\hat{f})=\lceil m / 2\rceil$, we have

$$
\text { dnf_size }(\hat{g})=s\left(\left\lceil\frac{m}{2}\right\rceil+1\right) \geq s\left(\frac{n-3}{4}+1\right)
$$

where $s$ is the number of assignments $x$ for which $\hat{f}(x)=*$.
We will upper bound $\operatorname{ess}{ }^{d}(\hat{g})$ by dividing the truepoints of $\hat{g}$ into two disjoint sets and upper-bounding the size of a maximum independent set of truepoints in each. (Recall that two truepoints of $\hat{g}$ are independent if they do not satisfy a common implicant of $\hat{g}$.)

Set 1: The set of all truepoints of $\hat{g}$ whose $x$ component has the property $f(x)=*$.
Let $a_{1}$ be a maximum independent set of truepoints of $\hat{g}$ consisting only of points in this set. Consider two truepoints $p$ and $q$ in this set that have the same $x$ value. It follows that they share the same values for $y_{1}$ and $y_{2}$. Let $t$ be the
term containing all variables $x_{i}$, and exactly one of the two $y_{j}$ variables, such that each $x_{i}$ appears without negation if it set to 1 by $p$ and $q$, and with negation otherwise, and $y_{j}$ is set to 1 by both $p$ and $q$. Clearly, $t$ is an implicant of $\hat{g}$ by the definition of $\hat{g}$, and clearly $t$ covers both $p$ and $q$. It follows that $p$ and $q$ are not independent.
Because any two truepoints in this set with the same $x$ value are not independent, $\left|a_{1}\right|$ cannot exceed the number of different $x$ assignments. There are $s$ assignments such that $\hat{f}(x)=*$, so $\left|a_{1}\right| \leq s$.
Set 2: The set of all truepoints of $\hat{g}$ whose $x$ component has the property $\hat{f}(x)=1$.
Let $a_{2}$ be a maximum independent set consisting only of points in this set. Consider any two truepoints $p$ and $q$ in this set that contain the same assignment for $z$. We can construct a term $t$ of the form $w y_{1} y_{2} \tilde{z}$ such that $w$ contains exactly $m-2$ of the $x_{i}$ variables that are set to 1 by both $p$ and $q$, and all $z_{i}$ s that are set to 1 by $p$ and $q$ appear in $\widetilde{z}$ without negation, and all other $z_{i} s$ appear with negation. It is clear that $t$ is an implicant of $\hat{g}$ and that $t$ covers both $p$ and $q$. Once again, it follows that $p$ and $q$ are not independent truepoints of $g$.
Because any two truepoints in this set with the same $z$ value are not independent, $\left|a_{2}\right|$ cannot exceed the number of different $z$ assignments. There are $s$ assignments to $z$ such that $z \in S$, so $\left|a_{2}\right| \leq s$.
Since a maximum independent set of truepoints of $\hat{g}$ can be partitioned into an independent set of points from the first set, and an independent set of points from the second set, it immediately follows that ${ }^{3}$

$$
\operatorname{ess}^{d}(\hat{g}) \leq\left|a_{1}\right|+\left|a_{2}\right| \leq s+s=2 s
$$

Hence, the ratio between the DNF size and ess(g) size is:

$$
\frac{s\left(\frac{n-3}{4}+1\right)}{2 s} \geq \frac{n+1}{8}=\Theta(n)
$$

Note that the above function gives a class of functions satisfying the conditions of Proposition 1 , for $\gamma=1$.
Corollary 1. There exists a function $f$ such that $\frac{c n f \_ \text {size }(f)}{\text { ess }(f)} \geq c n f$ _size $(f)^{\epsilon}$ for an $\epsilon \geq 0$.
Proof. In the previous construction, $\hat{f}(x)=*$ for exactly $\binom{m}{2}+1$ points, yielding $s=\Theta\left(n^{2}\right)$. Hence, the DNF size is $\Theta\left(m^{3}\right)$, making the ratio between $d n f \_\operatorname{size}(\hat{g})$ and $\operatorname{ess}^{d}(\hat{g})$ at least $\Theta\left(d n f \_\operatorname{size}(\hat{g})^{\frac{1}{3}}\right)$. The CNF result follows by duality.

### 4.2. Constructing a function with an exponential gap

Theorem 2. There exists a function $f$ on $n$ variables such that $\frac{c n f \_ \text {size }(f)}{\text { ess }(f)} \geq 2^{\Theta(n)}$.
Proof. As before, we will reduce a set covering instance to a DNF-minimization problem involving a partial Boolean function $f$. However, here we will rely on a more general version of Gimpel's reduction, due to Allender et al., described in the following lemma.

Lemma 1 ([1]). Let $s=\left\{S_{1}, \ldots, S_{p}\right\}$ be a set of subsets of ground set $A=\left\{e_{1}, \ldots, e_{m}\right\}$. Let $t>0$ and let $V=\left\{v^{i}: i \in\right.$ $\{1, \ldots, m\}\}$ and $W=\left\{w^{j}: j \in\{1, \ldots, p\}\right\}$ be sets of vectors from $\{0,1\}^{t}$ such that for all $j \in\{1, \ldots, p\}$ and $i \in\{1, \ldots, m\}$,

$$
e_{i} \in S_{j} \quad \text { iff } \quad v^{i} \geq w^{j}
$$

Let $f:\{0,1\}^{t} \rightarrow\{0,1, *\}$ be the partial function such that

$$
f(x)= \begin{cases}1 & \text { if } x \in V \\ * & \text { if } x \geq w \text { for some } w \in W \text { and } x \notin V \\ 0 & \text { otherwise }\end{cases}
$$

Then $s$ has a minimum cover of size $k$ iff dnf_size $(f)=k$.
(Note that the construction in the above lemma is equivalent to Gimpel's if we take $t=m, V=\left\{v \in\{0,1\}^{m} \mid v\right.$ contains exactly $m-1$ ones $\}$, and $W=\left\{x_{S} \mid S \in \delta\right\}$, where $x_{S}$ denotes the assignment in $\{0,1\}^{m}$ where $x_{i}=0$ iff $e_{i} \in S$.)

As before, we use the simple 2 -uniform set covering instance over $m$ elements where $s$ consists of all subsets of two of those elements. The next step is to construct sets $V$ and $W$ satisfying the properties in the above lemma for this set covering instance. To do this, we use a randomized construction of Allender et al. that generates sets $V$ and $W$ from an $r$-uniform set-covering instance, for any $r>0$. This randomized construction appears in the Appendix of [1], and is described in the following lemma.

[^1]Lemma 2. Let $r>0$ and let $s=\left\{S_{1}, \ldots, S_{p}\right\}$ be a set of subsets of $\left\{e_{1}, \ldots, e_{m}\right\}$, where each $S_{i}$ contains exactly $r$ elements. Let $t \geq 3 r(1+\ln (p m))$. Let $V=\left\{v^{1}, \ldots, v^{m}\right\}$ be a set of $m$ vectors of length $t$, where each $v^{i} \in V$ is produced by randomly and independently setting each bit of $v^{i}$ to 0 with probability $1 / r$. Let $W=\left\{w^{1}, \ldots, w^{p}\right\}$, where each $w^{j}=$ the bitwise AND of all $v^{i}$ such that $e_{i} \in S_{j}$. Then, the following holds with probability greater than $1 / 2$ : For all $j \in\{1, \ldots, p\}$ and $i \in\{1, \ldots, m\}, e_{i} \in S_{j}$ iff $v^{i} \geq w^{j}$.

By Lemma 2, there exist sets $V$ and $W$, each consisting of vectors of length $6\left(1+\ln \left(m^{2}(m-2) / 2\right)\right)=O(\log m)$, satisfying the conditions of Lemma 1 for our simple 2-uniform set covering instance. Let $\tilde{f}$ be the partial function on $O(\log m)$ variables obtained by using these $V$ and $W$ in the definition of $f$ in Lemma 1.

The DNF-size of $\tilde{f}$ is the size of the smallest set cover, which is $\lceil m / 2\rceil$, and the number of variables $n=\Theta(\log m)$; hence the DNF size is $2^{\Theta(n)}$.

We can convert the partial function $\tilde{f}(x)$ to a total function $\tilde{g}(x)$ just as done in the previous section. The arguments regarding DNF-size and $\operatorname{ess^{d}}(\tilde{g})$ remain the same. Hence, the DNF-size is now $s\left(2^{\Theta(n)}+1\right)$, and ess ${ }^{d}(\tilde{g})$ is again at most $2 s$.

The ratio between the DNF-size and $\operatorname{ess}^{d}(\tilde{g})$ is therefore at least $2^{\Theta(n)}$. Once again, the CNF result follows.

### 4.3. The quantity $\operatorname{ess}_{k}(f)$

We say that a set $S$ of falsepoints (truepoints) of $f$ is a " $k$-independent set" if no $k$ of the falsepoints (truepoints) of $f$ can be covered by the same implicate (implicant) of $f$.

We define $\operatorname{ess}_{k}(f)$ to be the size of the largest $k$-independent set of falsepoints of $f$, and $\operatorname{ess}_{k}^{d}(f)$ to be the size of the largest $k$-independent set of truepoints of $f$.

If $S$ is a $k$-independent set of falsepoints of $f$, then each implicate of $f$ can cover at most $k-1$ falsepoints in $S$. We thus have the following lower-bound on cnf_size(f): cnf_size $(f) \geq \frac{e s s_{k}(f)}{k-1}$.

Like ess(f), this lower bound is not tight.
Theorem 3. For any arbitrary $2 \leq k \leq h(n)$, where $h(n)=\Theta(n)$, there exists a function $f$ on $n$ variables, such that the gap between cnf_size(f) and $\frac{\operatorname{ess}_{k}(f)}{k-1}$ is at least $2^{\Theta\left(\frac{n}{k}\right)}$.

Proof. Consider the $k$-uniform set cover instance consisting of all subsets of $\left\{e_{1}, \ldots, e_{m}\right\}$ of size $k$. Construct $V$ and $W$ $\underset{\sim}{f}$ randomly using the construction from the Appendix of [1] described in Lemma 2, and define a corresponding partial function $\tilde{f}$, as in Lemma 1. Note that according to the definition of $\tilde{f}$, there can be no $k v^{i}$ for any $k$ values of $i \in\{1, \ldots, m\}$, such that all $v^{i} \geq w^{j}$ for some $j \in\{1, \ldots, p\}$. The maximum size $k$-independent set of truepoints of $\tilde{f}$ consists of $k-1$ truepoints.

We can convert the partial function $\tilde{f}$ to a total function $\tilde{g}$ according to the construction detailed in Section 4.1. Once again, we introduce $s$ new truepoints such that $\tilde{f}(x)=*$, yielding a maximum of $s$ pairwise independent truepoints. Any set of $k$ truepoints in $\tilde{g}$ that correspond to the same truepoint in $\tilde{f}$ must violate $k$-independence. Hence, the largest $k$-independent set of these points can contain a maximum of $s(k-1)$ points.

Any set of ground elements (i.e. truepoints of $\tilde{f}$ ) containing $k$ or more elements is not $k$-independent. Since $\tilde{g}$ has $s$ truepoints for each truepoint in $\tilde{f}$, and the points corresponding to the $s$ assignments to $z$ are all independent, the largest independent set for points of this type is of size no greater than $s(k-1)$. Since these two types of truepoints are disjoint, $e s s_{k}^{d}(\tilde{g}) \leq 2 s(k-1)$.

Since $\operatorname{ess}_{k}^{d}(\tilde{g}) / k-1 \leq 2 s(k-1) /(k-1)=2 s$, the ratio between $\operatorname{ess}_{k}^{d}(\tilde{g}) / k-1$ and dnf_size $(\tilde{g})$ is

$$
\frac{s\left(2^{\Theta\left(\frac{n}{k}\right)}+1\right)}{2 s} \geq 2^{\Theta\left(\frac{n}{k}\right)}
$$

The CNF result clearly follows.

## 5. Size of the gap for Horn functions

Because Horn-CNFs contain at most one unnegated variable per clause, they can be expressed as implications; e.g. $\neg a \vee \neg b \vee c$ is equivalent to $a b \rightarrow c$. Moreover, a conjunction of several clauses that have the same antecedent can be represented as a single meta-clause, where the antecedent is the antecedent common to all the clauses and the consequent is comprised of a conjunction of all the consequents, e.g. $(a \rightarrow b) \wedge(a \rightarrow c)$ can be represented as $a \rightarrow(b \wedge c)$.

### 5.1. Bounds on the ratio between cnf_size(f) and ess(f)

Angluin et al. [2] presented an algorithm (henceforth: the AFP algorithm) to learn Horn-CNFs, where the output is a series of meta-clauses. It can be proven $[3,4]$ that the output of the algorithm is of minimum implication size
(henceforth: $\min \_\operatorname{imp}(f)$ )-that is, it contains the fewest number of meta-clauses needed to represent function $f$. Each metaclause can be a conjunction of at most $n$ clauses; hence, each implication is equivalent to the conjunction of at most $n$ clauses. Therefore,

$$
c n f \_ \text {size }(f) \leq n \times \text { min_imp }(f)
$$

The learning algorithm maintains a list of negative and positive examples (falsepoints and truepoints of the Horn function, respectively), containing at most min_imp( $f$ ) examples of each.

Lemma 3. The set of negative examples maintained by the AFP algorithm is an independent set.
Proof. This proof relies heavily on [4]; see there for further details.
Let us consider any two negative examples $n_{i}$ and $n_{j}$ maintained by the algorithm. Without loss of generality, assume $i<j$. Then, Arias and Balcázar prove (Lemma 14 in [4]) that there exists a positive example $z$ such that $n_{i} \wedge n_{j} \leq z \leq n_{j}$. Clearly, $z$ separates $n_{i}$ and $n_{j}$. Hence, $n_{i}$ and $n_{j}$ are independent.

Theorem 4. For any Horn function $f, \frac{c n f \_ \text {size(f) }}{\text { ess }(f)} \leq n$.
Proof. For any Horn function $f$, there exists a set of negative examples of size at most min_imp(f), and these examples are all independent. Hence, $\operatorname{ess}(f) \geq \min \_i m p(f)$. We have already stated that cnf_size $(f) \leq n \times$ min_imp $(f)$ for this function.

Hence, cnf_size $(f) \leq n \times \operatorname{ess}(f)$.
Moreover, since Lemma 3 holds for general Horn functions in addition to pure Horn [4], this bound holds for all Horn functions.

### 5.2. Constructing a Horn function with a large gap between ess(f) and cnf_size(f)

Theorem 5. There exists a pure Horn function $f$ on $n$ variables such that $\frac{c n f \_ \text {size(f) }}{\text { ess } f)}=\Omega(\sqrt{n})$.
Proof. Consider the 2-uniform set covering instance over $k$ elements consisting of all subsets of two elements. We can construct a pure Horn formula $\varphi$ corresponding to this set covering according to the construction in [8], with modifications based on [5].

The formula $\varphi$ will contain 3 types of variables:

- Element variables: There is a variable $x$ for each of the $k$ elements.
- Set variables: There is a variable $s$ for each of the $\binom{k}{2}$ subsets.
- Amplification variables: There are $t$ variables $z_{1} \cdots z_{t}$.

The clauses in $\varphi$ are precisely the clauses in the following 3 groups:

- Witness clauses: There is a clause $s_{j} \rightarrow x_{i}$ for each subset and for each element that the subset covers. There are $2\binom{k}{2}$ such clauses.
- Feedback clauses: There is a clause $x_{1} \cdots x_{k} \rightarrow s_{j}$ for each subset. There are $\binom{k}{2}$ such clauses.
- Amplification clauses: There is a clause $z_{h} \rightarrow s_{j}$ for every $h \in\{1 \cdots t\}$ and for every subset. There are $t\binom{k}{2}$ such clauses.

It follows from [8] that any minimum CNF for this function must contain all witness and feedback clauses, along with tc amplification clauses, where $c$ is the size of the smallest set cover.

This particular function $f$ has a minimum set cover of size $k / 2$; hence, cnf_size $(f)=2\binom{k}{2}+\binom{k}{2}+t(k / 2)$.
We will upper bound $\operatorname{ess}(f)$ by dividing the falsepoints of $f$ into three disjoint sets and bounding the size of the maximum independent set for each.

Set 1: The set of all falsepoints of $f$ that contain at least one $x_{i}=0$ for $i \in\{1, \ldots, k\}$ and some $s_{j}=1$ for a subset $s_{j}$ that covers $x_{i}$.
Let $a_{1}$ be an independent set of $f$ consisting of points in this set. These points can be covered by implicates of the form $s_{j} \rightarrow x_{i}$, of which there are $2\binom{k}{2}$. If two points in the set both have $x_{i}=0$ and $s_{j}=1$ for a subset $s_{j}$ that covers $x_{i}$, then they are both covered by $s_{j} \rightarrow x_{i}$ and are not independent. Hence $a_{1}$ can contain no more than $2\binom{k}{2}$ points.
Set 2: The set of all falsepoints that are not in the first set, have $x_{i}=1$ for all $i \in\{1, \ldots, k\}$, and at least one $s_{j}=0$ for some $j \in\left\{1, \ldots,\binom{k}{2}\right\}$.

Let $a_{2}$ be the largest independent set of $f$ consisting of points in this set. These points can be covered by implicates of the form $x_{1} \cdots x_{k} \rightarrow s_{j}$. There are $\binom{k}{2}$ such implicates. Hence, by the same argument as above, $a_{2}$ can contain no more than $\binom{k}{2}$ points.
Set 3: The set of all falsepoints that are not in the first two sets, and therefore have $z_{h}=1$ for some $h \in\{1, \ldots, t\}, x_{i}=0$ for some $i \in\{1, \ldots, k\}$, and $y_{j}=0$ for all subsets $y_{j}$ covering $x_{i}$.
Let $a_{3}$ be an independent set of $f$ consisting of points in this set. Consider a falsepoint $p$ in this set where $x_{i}=0$ for at least one $i \in\{1, \ldots, k\}$. If $p$ contained a $y_{j}=1$ such that the subset $y_{j} \operatorname{covers} x_{i}$, that point would be a point in the first set. Hence, the only points of this form in this set have $y_{j}=0$ for all $k-1$ subsets $y_{j}$ that cover $x_{i}$.
Now consider another falsepoint $q$ in this set, where $x_{a}=0$ for at least one $a \in\{1, \ldots, k\}$. Once again, the only points in this set must set $y_{b}=0$ for all $k-1$ subsets $y_{b}$ that cover $x_{a}$.
Because the set covering problem included a set for each pair of $x_{i}$ points, there exists some $y_{j}$ that covers both $x_{i}$ and $x_{a}$. By the previous argument, that $y_{j}$ is set to 0 in all assignments that set $x_{i}$ or $x_{a}=0$. If for some $h, z_{h}=1$ in both $p$ and $q$, then $p$ and $q$ can be covered by the implicate $z_{h} \rightarrow y_{j}$. Hence, points $p$ and $q$ are not independent.
In fact, any two falsepoints chosen that are not in the first set and contain $z_{h}=1$ for the same $h$ and at least one $x_{i}=0$ are not independent. Because there are $t$ values of $h$, size at most $t$.
The largest independent set for all falsepoints cannot exceed the sum of the independent sets for these three disjoint sets, hence

$$
\operatorname{ess}(f) \leq\left|a_{1}\right|+\left|a_{2}\right|+\left|a_{3}\right| \leq 2\binom{k}{2}+\binom{k}{2}+t
$$

The gap between cnf_size(f) and

$$
\operatorname{ess}(f)=\frac{c n f \_ \text {size }(f)}{\operatorname{ess}(f)} \geq \frac{3\binom{k}{2}+t(k / 2)}{3\binom{k}{2}+t}
$$

Let us set $t=3\binom{k}{2}$. The difference is now:

$$
\frac{c n f_{\_} \operatorname{size}(f)}{\operatorname{ess}(f)} \geq \frac{t(1+k / 2)}{2 t}=\Theta(k)
$$

We have $k$ element variables, $\binom{k}{2}$ set variables, and $3\binom{k}{2}$ amplification variables, yielding $n=\Theta\left(k^{2}\right)$ variables in total. The ratio between cnf_size $(f)$ and $\operatorname{ess}(f)$ is therefore $\Theta(\sqrt{n})$.

We earlier posited that if $\Sigma_{p}^{2} \neq c o-N P$, there exists an infinite set of functions for which $\frac{c n f-\operatorname{size}(f)}{\text { ess }(f)} \geq c n f \_$size $(f)^{\gamma}$ for some $\gamma>0$. We can now prove a stronger theorem:

Theorem 6. There exists an infinite set of Horn functions $f$ for which $\frac{c n f \_\operatorname{size}(f)}{\text { ess }(f)} \geq \mathrm{cnf}$ _size $(f)^{\gamma}$.
Proof. See construction above. Because cnf_size $(f)=\Theta\left(k^{3}\right), \frac{c n f \_\operatorname{size}(f)}{\text { ess }(f)}=\Theta\left(c n f \_\operatorname{size}(f)^{1 / 3}\right)$.

## Acknowledgments

This work was partially supported by the US Department of Education GAANN grant P200A090157, and by NSF Grant CCF-0917153.

## References

[1] E. Allender, L. Hellerstein, P. McCabe, T. Pitassi, M.E. Saks, Minimizing disjunctive normal form formulas and AC ${ }^{0}$ circuits given a truth table, SIAM Journal on Computing 38 (2008) 63-84.
[2] D. Angluin, M. Frazier, L. Pitt, Learning conjunctions of horn clauses, Machine Learning 9 (1992) 147-164.
[3] M. Arias, J.L. Balcázar, Query learning and certificates in lattices, in: Y. Freund, L. Györfi, G. Turán, T. Zeugmann (Eds.), Algorithmic Learning Theory, in: Lecture Notes in Computer Science, vol. 5254, Springer, Berlin, Heidelberg, 2008, pp. 303-315.
[4] M. Arias, J. Balcázar, Construction and learnability of canonical horn formulas, Machine Learning (2011) 1-25. http://dx.doi.org/10.1007/s 10994-011-5248-5.
[5] A. Bhattacharya, B. DasGupta, D. Mubayi, G. Turán, On approximate horn formula minimization, in: S. Abramsky, C. Gavoille, C. Kirchner, F. Meyer auf der Heide, P. Spirakis (Eds.), Automata, Languages and Programming, in: Lecture Notes in Computer Science, vol. 6198, Springer, Berlin, Heidelberg, 2010, pp. 438-450.
[6] O. Čepek, P. Kučera, P. Savický, Boolean Functions with a simple certificate for CNF complexity, Technical Report, Rutgers Center for Operations Research, 2010.
[7] O. Coudert, Two-level logic minimization: an overview, Integration, the VLSI Journal (1994).
[8] Y. Crama, P.L. Hammer (Eds.), Boolean Functions: Theory, Algorithms, and Applications, Cambridge University Press, 2011.
[9] S.L.A. Czort, The complexity of minimizing disjunctive normal form formulas, Master's Thesis, University of Aarhus, Aarhus, Denmark, 1999.
[10] J. Gimpel, Method of producing a Boolean function having an arbitrarily prescribed prime implicant table, IEEE Transactions on Computers (1965).
[11] S.J. Russell, P. Norvig, Artificial intelligence: a modern approach, Pearson Education (2003).
[12] C. Umans, Hardness of approximating $\Sigma_{2}^{p}$ minimization problems, in: Proc. IEEE Symposium on Foundations of Computer Science, pp. $465-474$.
[13] C. Umans, The minimum equivalent DNF problem and shortest implicants, in: IEEE Symposium on Foundations of Computer Science, pp. 556-563.


[^0]:    * Corresponding author. Tel.: +1 347587 3112; fax: +1 5304833112.

    E-mail addresses: hstein@poly.edu (L. Hellerstein), dkletenik@cis.poly.edu (D. Kletenik).
    1 This definition immediately follows from Corollary 3.2 of Čepek et al. [6].
    2 Their function is actually defined in terms of two parameters $n_{1}$ and $n_{2}$. Setting them to maximize the multiplicative gap between ess( $f$ ) and $c n f$ _size ( $f$ ), as a function of the number of variables $n$, yields a gap of $\operatorname{size} \Theta(\log n)$.

[^1]:    3 It can actually be proved that in fact, $\operatorname{ess}^{d}(\hat{g})=2 s$, but details of this proof are omitted.

