Von Neumann's Minimax Theorem in the Gentry Wichs Separation

Christopher Smith

Last Updated: November 5, 2025

1 Overview

The Gentry Wichs separation [GW11] is a fundamental barrier in theoretical cryptography stating that the adaptive soundness of any SNARG cannot be proved with a black-box reduction to any falsifiable assumption. It should be mentioned that recent results [WW24,WZ24,WW25] circumvent this barrier through clever uses of complexity leveraging (and pay for this in CRS size). Barrier busting aside, some of the techniques used in [GW11] are quite interesting. In particular, the proof of their key "indistinguishability with auxiliary information" lemma relies on an application of von Neumann's minimax theorem. The aim of this note is to introduce this minimax theorem and unpack its usage within [GW11].

2 Minimax Theorems

See the Wikipedia page on minimax theorems [Wik25b] for a solid introduction to the topic. For more in-depth resources, see Sion [Sio58] and Kjeldsen [Kje01]. In short, a minimax theorem is a theorem claiming something of the form

$$\max_{x\in X}\min_{y\in Y} f(x,y) = \min_{y\in Y}\max_{x\in X} f(x,y)$$

under specific conditions on X, Y, and f. Any point (x,y) at which the equality holds is often called a *saddle point*. John von Neumann [vN28] is credited with the first of these theorems, which is stated as follows.

Theorem 1 (von Neumann minimax [Wik25b])**.** *Let $X = \{(x_1,...,x_n) \in [0,1]^n \mid \sum x_i = 1\}$ and $Y = \{(y_1,...,y_m) \in [0,1]^m \mid \sum y_i = 1\}$ be standard simplexes, and let $f(x,y)$ be a linear function in both of its arguments (that is, f is bilinear), and can therefore be written as $f(x,y) = x^\top Ay$ for a finite matrix $A \in \mathbb{R}^{n\times m}$, or equivalently as $f(x,y) = \sum_{i=1}^n \sum_{j=1}^m A_{ij}x_iy_j$. Then*

$$\max_{x\in X}\min_{y\in Y} f(x,y) = \min_{y\in Y}\max_{x\in X} f(x,y)$$

Type Inference for f. Note that f technically cannot be bilinear over $X \times Y$ as X, Y are not proper vector spaces. It follows that f must at least be bilinear over some pair of vector spaces $\hat{X} \supset X, \hat{Y} \supset Y$. From the characterization in Theorem 1 of f as $\sum_{i=1}^n \sum_{j=1}^m A_{ij}x_iy_j$ where A_{ij} are entries in a finite matrix $A \in \mathbb{R}^{n\times m}$, we can deduce that the scalar field of f is $\mathbb{R}$, and that $\hat{X}, \hat{Y}$ must be finite dimensional. Specifically, $\dim \hat{X} \le n$ and $\dim \hat{Y} \le m$. But notice that any vector space containing X must necessarily contain $\mathsf{span}(X)$, and $\mathsf{span}(X) = \mathbb{R}^n$ since the n standard

basis vectors for for $\mathbb{R}^n$ already belong to X. Thus, $\hat{X}$ must be isomorphic to $\mathbb{R}^n$, and $\hat{Y}$ must be isomorphic to $\mathbb{R}^m$. All this to conclude that we can safely think of f as a bilinear function $\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ (*i.e.*, a bilinear form, depending on whether you insist that $n = m$).

Bilinear vs. continuous quasi-concave-convex I feel compelled to mention that, as explained in [Kje01], von Neumann actually proved his theorem for the more general case of a continuous function $f : X \times Y \to R$ that is quasiconcave in in X and quasiconvex in Y (Definition 5 implies X and Y are convex subsets of a real vector space). Perhaps to nobody's surprise, the bilinear function $f(x, y) = \sum_i \sum_j A_{ij} x_i y_j$ of Theorem 1 satisfies these conditions. Let us prove this anyway.

Claim 2. *Let $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ be bilinear. Then f is continuous, quasiconcave in its first argument, and quasiconvex in its second argument.*

Proof. We first prove continuity. We do this by showing that f is separately continuous (*i.e.*, $f(x, \cdot)$ and $f(\cdot, y)$ are continuous $\forall x, y$), and then appealing to the fact [Wik25a] that a separately continuous bilinear map $f : X \times Y \to Z$ is continuous if X is a Baire space and Y is metrizable. Indeed, since R^n is a complete metric space it is both Baire and metrizable, so it remains to be seen that f is separately continuous. By symmetry it is enough to show that $f(\cdot, y) : \mathbb{R}^n \to \mathbb{R}$ is continuous. Because $\mathbb{R}^n$ is a normed and finite dimensional space, by Theorem 2.7-8 of [Kre91], it is bounded. By Theorem 2.7-9 of [Kre91], since $f(\cdot, y)$ is a bounded linear operator and $\mathbb{R}^n, \mathbb{R}$ are normed spaces, $f(\cdot, y)$ is continuous.

Next we show $f(\cdot, y)$ is quasiconcave for all y. Let $x_1, x_2 \in \mathbb{R}^n$, $t \in [0, 1]$ be arbitrary. We have:

$$f(tx_1 + (1-t)x_2, y) = tf(x_1, y) + (1-t)f(x_2, y) \geq \min\{f(x_1, y), f(x_2, y)\}$$

Similarly we show quasiconvexity for $f(x, \cdot)$ for all x. Let $y_1, y_2 \in \mathbb{R}^m$, $t \in [0, 1]$. We have:

$$f(x, ty_1 + (1-t)y_2) = tf(x, y_1) + (1-t)f(x, y_2) \leq \max\{f(x, y_1), f(x, y_2)\}$$

□

3 Minimax in Gentry Wichs

In order to see exactly how minimax is applied in [GW11], we introduce relevant notation from the paper. Let $\mathsf{size}(m)$ denote the set of all circuits of size m, and $\mathsf{dist}(m)$ denote the set of all distributions over $\mathsf{size}(m)$. So for example (specifically the example from [GW11]), if $s^*(\cdot)$ is some polynomial, then $\mathsf{dist}(s^*(n) + 1)$ is the set of all distributions over circuits of size $s^*(n) + 1$. Further, fix some distribution ensembles $\mathcal{L}_n$ and $\overline{\mathcal{L}}_n$ over a language L and its complement $\bar{L}$ (there are further conditions on L in the paper but they are not important if we just want to see how minimax is applied), and fix some joint distribution $\mathcal{L}_n^*$ over tuples (x, π) such that $x \leftarrow \mathcal{L}_n$, and let $\mathsf{dist}(\overline{\mathcal{L}}_n)$ be the set of all joint-distributions on tuples $(\bar{x}, \bar{\pi})$ with $\bar{x} \leftarrow \overline{\mathcal{L}}_n$.

Additionally, for purposes of this note let $\bar{\ell}^*(n) \coloneqq |\bar{x}| + |\bar{\pi}|$ for $(\bar{x}, \bar{\pi}) \leftarrow \overline{\mathcal{L}}_n^* \in \mathsf{dist}(\overline{\mathcal{L}}_n)$. Then $\overline{\mathcal{L}}_n^*$ takes values in $\bar{L} \cap \{0,1\}^{\bar{\ell}^*(n)}$, and we can let $M \coloneqq |\bar{L} \cap \{0,1\}^{\bar{\ell}^*(n)}|$ denote the size of the support for any $\overline{\mathcal{L}}_n^* \in \mathsf{dist}(\overline{\mathcal{L}}_n)$. Similarly, let $N \coloneqq |\mathsf{size}(s^*(n) + 1)|$ denote the size of the support for any $\mathbb{D}_n \in \mathsf{dist}(s^*(n) + 1)$.

By von Neumann's minimax theorem, the proof of Lemma 3.1 in the paper claims the following equality:

$$\min_{\overline{\mathcal{L}}_n^* \in \mathsf{dist}(\overline{\mathcal{L}}_n)} \max_{\mathbb{D}_n \in \mathsf{dist}(s^*(n)+1)} \mathop{\mathbb{E}}_{\substack{(\bar{x},\bar{\pi}) \leftarrow \overline{\mathcal{L}}_n^* \\ D_n \leftarrow \mathbb{D}_n}} \left[D_n(\bar{x},\bar{\pi}) - \Pr_{(x,\pi)\leftarrow\mathcal{L}_n^*}[D_n(x,\pi)=1] \right]$$

$$= \max_{\mathbb{D}_n \in \mathsf{dist}(s^*(n)+1)} \min_{\overline{\mathcal{L}}_n^* \in \mathsf{dist}(\overline{\mathcal{L}}_n)} \mathop{\mathbb{E}}_{\substack{(\bar{x},\bar{\pi}) \leftarrow \overline{\mathcal{L}}_n^* \\ D_n \leftarrow \mathbb{D}_n}} \left[D_n(\bar{x},\bar{\pi}) - \Pr_{(x,\pi)\leftarrow\mathcal{L}_n^*}[D_n(x,\pi)=1] \right]$$

In light of Sec. 2, in order to invoke Theorem 1 in this way, it must be the case that $\mathsf{dist}(\overline{\mathcal{L}}_n)$ and $\mathsf{dist}(s^*(n)+1)$ are standard simplexes, and that the above expectation is bilinear.

Indeed, $\mathsf{dist}(\overline{\mathcal{L}}_n)$ is the set of all finite probability distributions supported over $L \cap \{0,1\}^{\bar{\ell}^*(n)}$, and each $\overline{\mathcal{L}}_n^* \in \mathsf{dist}(\overline{\mathcal{L}}_n)$ can be represented by a vector $(p_1,...,p_M)$, where p_i is the probability that $\overline{\mathcal{L}}_n^*$ takes the value of the i-th tuple $(\bar{x},\bar{\pi})_i \in \bar{L} \cap \{0,1\}^{\bar{\ell}^*(n)}$. Thus, $\mathsf{dist}(\overline{\mathcal{L}}_n)$ is the standard $(M-1)$-simplex. Similarly, $\mathsf{dist}(s^*(n)+1)$ is the standard $(N-1)$-simplex, and any $\mathbb{D}_n \in \mathsf{dist}(s^*(n)+1)$ can be represented by a vector $(q_1,...,q_N)$ where q_j is the probability that $\mathbb{D}_n$ takes the value of the j-th circuit $D_n^{(j)} \in \mathsf{size}(s^*(n)+1)$.

Now we show that the expectation expression above is bilinear. Begin by defining the function $g : (\bar{L} \cap \{0,1\}^{\bar{\ell}^*(n)}) \times \mathsf{size}(s^*(n)+1) \to \mathbb{R}$ as follows:

$$g((\bar{x},\bar{\pi}),D_n) \coloneqq D_n(\bar{x},\bar{\pi}) - \Pr_{(x,\pi)\leftarrow\mathcal{L}_n^*}[D_n(x,\pi)=1]$$

Because we can index the values in the domain of g with elements of $[M]\times[N]$, we can equivalently define $g : [M]\times[N] \to \mathbb{R}$ as:

$$g(i,j) \coloneqq D_n^{(j)}((\bar{x},\bar{\pi})_i) - \Pr_{(x,\pi)\leftarrow\mathcal{L}_n^*}\left[D_n^{(j)}(x,\pi)=1\right]$$

Now let $f|_{\mathsf{simplex}} : \mathsf{dist}(\overline{\mathcal{L}}_n) \times \mathsf{dist}(s^*(n)+1) \to \mathbb{R}$ be given by

$$\begin{aligned} f|_{\mathsf{simplex}}(\overline{\mathcal{L}}_n^*, \mathbb{D}_n) &\coloneqq \mathop{\mathbb{E}}_{\substack{(\bar{x},\bar{\pi}) \leftarrow \overline{\mathcal{L}}_n^* \\ D_n \leftarrow \mathbb{D}_n}} \left[g(\overline{\mathcal{L}}_n^*, \mathbb{D}_n) \right] \\ &= \sum_{(\bar{x},\bar{\pi})\in\overline{\mathcal{L}}_n^*} \sum_{D_n \in \mathbb{D}_n} g((\bar{x},\bar{\pi}),D_n) \Pr\left[\overline{\mathcal{L}}_n^* = (\bar{x},\bar{\pi})\right] \Pr\left[\mathbb{D}_n = D_n\right] \\ &= \sum_{i\in[M]} \sum_{j\in[N]} g(i,j) p_i q_j \end{aligned}$$

Written in this way, it is clear that the extension f of $f|_{\mathsf{simplex}}$ to $\mathbb{R}^N \times \mathbb{R}^M \to \mathbb{R}$ is bilinear, and therefore the use of the minimax theorem is justified.

A Supplementary Definitions

Definition 3 (Standard Simplex [Wik25d])**.** *The standard n-simplex (or unit n-simplex) is the subset of $\mathbb{R}^{n+1}$ given by*

$$\Delta^n = \left\{ (t_0,...,t_n) \in \mathbb{R}^{n+1} \;\middle|\; \sum_{i=0}^{n} t_i = 1 \text{ and } t_i \geq 0 \text{ for } i = 0,...,n \right\}$$

Definition 4 (Bilinear Map)**.** *Let X, Y, Z be three vector spaces over the same base field $\mathbb{F}$. A bilinear map is a function $f : X \times Y \to Z$ such that for all $y \in Y$, the map $f(\cdot, y) : X \to Z$ is linear, and for all $x \in X$, the map $f(x, \cdot) : Y \to Z$ is linear. If $Z = \mathbb{F}$, then f is a bilinear form.*

Definition 5 (Quasiconvex function [Wik25c])**.** *A function $f : S \to \mathbb{R}$ defined on a convex subset S of a real vector space is quasiconvex if for all $x, y \in S$ and $t \in [0, 1]$ we have*

$$f(tx + (1 - t)y) \leq \max\{f(x), f(y)\}$$

Alternatively, f is quasiconcave if $(-f)$ is quasiconvex, and in this case we have:

$$f(tx + (1 - t)y) \geq \min\{f(x), f(y)\}$$

References

[GW11] Craig Gentry and Daniel Wichs. Separating succinct non-interactive arguments from all falsifiable assumptions. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pages 99–108, 2011.

[Kje01] Tinne Hoff Kjeldsen. John von neumann's conception of the minimax theorem: A journey through different mathematical contexts. *Archive for history of exact sciences*, 56(1):39–68, 2001.

[Kre91] Erwin Kreyszig. *Introductory functional analysis with applications*. John Wiley & Sons, 1991.

[Sio58] Maurice Sion. On general minimax theorems. 1958.

[vN28] J v. Neumann. Zur theorie der gesellschaftsspiele. *Mathematische annalen*, 100(1):295–320, 1928.

[Wik25a] Wikipedia contributors. Bilinear map — Wikipedia, the free encyclopedia. `https://en.wikipedia.org/w/index.php?title=Bilinear_map&oldid=1311554362`, 2025. [Online; accessed 4-November-2025].

[Wik25b] Wikipedia contributors. Minimax theorem — Wikipedia, the free encyclopedia. `https://en.wikipedia.org/w/index.php?title=Minimax_theorem&oldid=1296417469`, 2025. [Online; accessed 29-October-2025].

[Wik25c] Wikipedia contributors. Quasiconvex function — Wikipedia, the free encyclopedia. `https://en.wikipedia.org/w/index.php?title=Quasiconvex_function&oldid=1303920671`, 2025. [Online; accessed 4-November-2025].

[Wik25d] Wikipedia contributors. Simplex — Wikipedia, the free encyclopedia. `https://en.wikipedia.org/w/index.php?title=Simplex&oldid=1314265483`, 2025. [Online; accessed 3-November-2025].

[WW24] Brent Waters and David J Wu. Adaptively-sound succinct arguments for np from indistinguishability obfuscation. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 387–398, 2024.

[WW25] Brent Waters and David J Wu. A pure indistinguishability obfuscation approach to adaptively-sound snargs for np. In *Annual International Cryptology Conference*, pages 292–326. Springer, 2025.

[WZ24] Brent Waters and Mark Zhandry. Adaptive security in snargs via io and lossy functions. In *Annual International Cryptology Conference*, pages 72–104. Springer, 2024.