RFocus: Beamforming Using Thousands of Passive Antennas

Venkat Arun and Hari Balakrishnan, *Massachusetts Institute of Technology*

https://www.usenix.org/conference/nsdi20/presentation/arun

RFocus: Beamforming Using Thousands of Passive Antennas

Venkat Arun and Hari Balakrishnan
MIT, CSAIL
{venkatar, hari}@csail.mit.edu

Abstract

To reduce transmit power, increase throughput, and improve range, radio systems benefit from the ability to direct more of the transmitted power toward the intended receiver. Many modern systems beamform with antenna arrays for this purpose. However, a radio's ability to direct its signal is fundamentally limited by its size. This limitation is acute on IoT and mobile devices, which are small and inexpensive, but even access points and base stations are typically constrained to a modest number of antennas.

To address this problem, we introduce *RFocus*, which moves beamforming functions from the radio endpoints to the environment. RFocus includes a two-dimensional surface with a rectangular array of simple RF switch elements. Each switch element either lets the signal through or reflects it. The surface does not emit any power of its own. The state of the elements is set by a software controller to maximize the signal strength at a receiver, with a novel optimization algorithm that uses signal strength measurements from the receiver. The RFocus surface can be manufactured as an inexpensive thin "wallpaper". In one floor of an office building (at MIT CSAIL), our prototype improves the median signal strength by $9.5\times$ and the median channel capacity by $2.0\times$.

1 Introduction

Many radio systems use directional or sectorized antennas and beamforming to improve the throughput or range of a wireless communication link. Beamforming ensures that a larger fraction of transmitted energy reaches the intended receiver, while reducing unintended interference. A radio with many antennas spread densely over a large area can fundamentally beamform better than a smaller radio [1, 4]. However, there are many practical challenges to making radio systems with large antenna arrays. First devices such as IoT sensors and handhelds must be small in size. Second, connecting each antenna in an array to full-fledged radio transmit/receive circuitry increases cost and power. Third, large, bulky systems are hard to deploy, even in infrastructure base stations or access points.

To address these challenges and achieve the equivalent of a large number of antennas, we propose *RFocus*. RFocus

includes a *software-controlled surface* placed in the environment, made up of thousands of simple switching elements. RFocus also has a *controller* that configures each element so that the surface, as a whole, directs a signal from a transmitter to a receiver without any signal amplification (i.e., no extra power) or knowledge of the locations of the transmitter or receiver. This approach moves the task of beamforming from the transmitter to the surface. Any device in the vicinity of the surface can reap the benefits of RFocus without itself being large or consuming additional energy.

The RFocus surface is made up of thousands of simple elements organized in a rectangular array. To reduce cost and energy, each element is a single simple 2-way RF switch. Each element in the RFocus surface can be in one of two states. When "on", the element reflects the signal; otherwise, the signal passes through. Each receiver periodically sends measurements of the received signal strength to the RFocus controller. The controller uses these to configure the elements on the RFocus surface to maximize the received signal strength. This optimization has two steps: a training phase, where the controller configures test states on the surface and monitors the changes in the reported measurements to collect data, and an optimized phase where the controller uses this data to set the on/off state for each element. RFocus can switch between optimized states to maximize signal strength between different pairs of endpoints.

RFocus can work both as a "mirror" or a "lens", with the controller choosing the right mode. That is, radio endpoints can be on the same side of the surface, or on opposite sides.

We have built an RFocus prototype with 3,200 inexpensive[1] antennas on a 6 square-meter surface. This configuration may well be the largest published number of antennas ever used for a single communication link.

The controller's optimization algorithm solves three key challenges. First, indoor environments exhibit complex multi-path, and the direct path may be blocked. Therefore the optimal configuration might not correspond to directing the signal along a single direction, but can depend on the environment. Second, the effect of an individual element on the channel is usually miniscule, and hence hard to measure (§6.2.3). Third, the phase of a channel is hard to

---

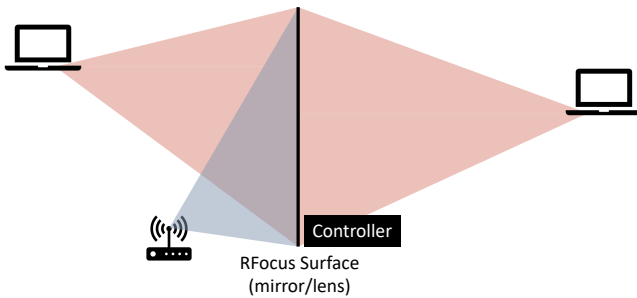[1] At scale, the cost of each element is a few cents.

Figure 1: RFocus is an electronically configurable mirror/lens that focuses the transmitter's signal on the receiver.
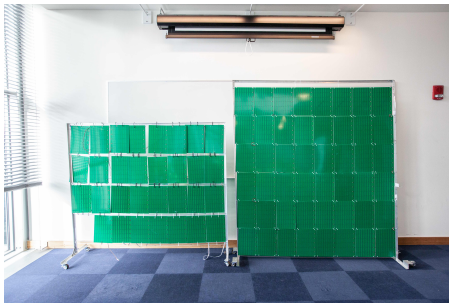

Figure 2: Our prototype of the RFocus surface.

measure when the signal is weak, due to carrier frequency offset and radio-clock jitter. And last but not least, commodity radio receivers usually don't report phase.

The contributions of this paper include:

1. The RFocus controller, incorporating two key ideas. First, it modulates all elements at once to boost the effect of the RFocus surface on the channel, hence making the change large enough to be measurable. Second, it relies only on signal strength measurements, sidestepping difficulties in measuring phase. We provide an analysis of the algorithm under some assumptions, that suggests that it finds a solution that is close to optimal.

2. Experiments, which show that in a typical indoor office environment, RFocus achieves a median $9.5\times$ improvement in signal strength and $2.0\times$ improvement in channel capacity. Moreover, RFocus is robust to element failure; even when one-third the elements fail, the relative performance improvement does not plummet to 0, but drops by 50%.

2 Related Work

Passive reflective surfaces to control incident radiation, called "meta-surfaces" have been studied extensively by applied physicists in the microwave, terahertz, and optical frequencies [2,

6–8, 13]. These works focus on hardware design to control the incident radiation efficiently. Some of these metasurfaces are configurable. Here, given the incident wave and desired reflected/transmitted wave direction(s), they develop algorithms to determine good configurations, for example to create programmable holograms. Recently there has been interest in using metasurfaces, called "reconfigurable intelligent surfaces", to improve communication channels. Most papers explore the idea with theory and simulation [12,17–19,28] except for a few papers providing small-scale experimental results [11, 15, 29] that evaluate channel improvement between one fixed pair of endpoints that are close to each other and to the surface.

In prior work, the method to determine a good configuration works by varying the state of the elements *one at a time* and measuring its effect on the channel. This approach is linear in the number of elements, but requires the receiver to be able to accurately measure the impact of the change in a single element. If the transmitter and receiver are very close to the surface, the effect is measurable, but not otherwise because the strength of the signal attenuates as at least $1/(d_s d_r)^2$, where d_s and d_r are the distances from the element to the sender and receiver, respectively. For a large surface with thousands of antennas, the effect becomes hard to measure because some elements will necessarily be at long distances from the transmitter or receiver §6.2.3. The reason is that the size of each element is between one-quarter to one-half of the wavelength of the signal (at 2.4 GHz, the wavelength is 15 cm), so thousands of elements require a surface area of a few square meters, and at a distance of a few meters, measuring a single-element change is inaccurate especially on commodity radio receivers. Our algorithm overcomes this problem and scales to larger surfaces by varying many elements at once and boosting the signal strength changes being measured.

LAIA [16, 32] is a recent project that also achieves gains at larger distances, but with a different setup and design from RFocus. LAIA helps endpoints whose line-of-sight path is blocked by a wall. LAIA elements collect radio energy from one side of the wall, phase-shift them, and take them to the other side of the wall via wires traversing holes in the wall. Because they go through the wall with wires, individual LAIA elements have a much larger impact on the channel than RFocus elements, allowing LAIA to function with fewer elements than RFocus, and without RFocus's optimization algorithm.

Another line of related work improves wireless coverage by analyzing the indoor space, and custom-designing a 3D reflector for that space. When 3D-printed and placed behind an access point (AP), the reflector reflects energy in specific directions to maximize signal strength at previously uncovered areas [5, 33]. Once deployed, this reflector has a low operational cost because it is just a passive metal-coated object. But a new reflector needs to be designed for each new location and whenever the space changes or the AP is moved or a new AP is added. Moreover, a single solution has to be designed for all pairs of endpoints, whereas RFocus dynamically and

automatically configures a new pattern for each pair.

RFocus can be thought of as a phased array [25, 26] that uses the air instead of transmission lines to distribute signals from the transmitter to the antennas. We believe this allows larger antenna arrays to be deployed more ubiquitously. Transmission lines tend to be expensive and bulky. Even PCB-trace transmission lines, which are easy to manufacture, depend on properties of the dielectric and cannot be manufactured as a paper-thin sheet. This is unlike the wires used in RFocus, which are digital and do not carry high-speed signals. Due to its simplicity, RFocus can be readily incorporated into the environment (§7). On the other hand, for the same number of antennas, RFocus achieves a lower gain than phased arrays because it doesn't harness all of the transmitted energy (§7).

Digital Light Processors (DLPs) and Digital Micromirror Devices (DMDs) [22, 23, 27] are micro-electro-mechanical systems (MEMS) devices that have a large array of small (microns or smaller) mirrors on a single silicon chip. The mirrors can be toggled between two angles, controlling which ones reflect light onto the object, and which reflect it elsewhere. They are used for steering lasers and in display projectors. RFocus could be thought of as a DMD for radio.

Range extenders increase signal strength at the receiver. However, by rebroadcasting signals, they increase interference and reduce transmission opportunities. By precisely focusing energy already available, RFocus decreases interference while increasing signal strength. That said, RFocus complements range extenders and both could be used together if necessary.

Reconfigurable antennas [10] and reflectarray antennas [14] have RF switches and phase shifters, which allow them to dynamically change their characteristics such as operating frequency, input impedance, and directionality. These approaches modify an antenna to improve characteristics. By contrast, RFocus leaves the transmit and receive antennas unmodified, instead modifying the environment to improve communication for all nearby devices. Additionally, like phased arrays, it is difficult to use wires to scale to a large number of radiating elements.

3 Background

3.1 System Model and Notation

In most environments, there are multiple paths between a sending and receiving antenna. For a narrow-band signal, the effect of each path can be represented by a complex number. The net effect of the channel is the sum of the effects of all the paths. A subset of the paths pass via each of the N elements on the RFocus surface. Denote the contribution of the elements by $c_1, ..., c_N$. We combine all the paths *not* going via RFocus into one complex number c_E (E for environment). The net channel, h, is equal to $c_E + \sum_{i=1}^{N} \alpha_i c_i$, where α_i represents the amplitude change and phase shift introduced by element i.

RFocus controls the channel, h, by adjusting α_i. c_E and c_i

are functions of the path lengths. α_i depends on three factors: the state of the ith element and its neighbors; the shape of the antennas composing the element, and the angles at which the path enters and leaves the ith element. Since the elements are passive and lack power, $|\alpha_i| \leq 1$.

Maximizing $|h|$ maximizes the received signal strength. The maximum possible value of $|h|$, the sum of several complex numbers, is the sum of the magnitudes of the summands:

$$|h_{\max}| = |c_E| + \sum_{i=1}^{N} |c_i| \qquad (1)$$

This maximum is achievable if we had full control over each α_i by setting $\alpha_i = \frac{c_E}{|c_E|} \frac{c_i^*}{|c_i|}$, where c_i^* denotes conjugation. Unfortunately, we do not have full control over α_i. Each element can only be on or off. If we assume that α_i is a function of only the ith element's state and not its neighbors (we validate this assumption with experiments in §6.2.1), then we can write the channel as

$$h = h_Z + \mathbf{h} \cdot \mathbf{b} = h_Z + \sum_{i=1}^{N} \mathbf{h}^{(i)} \mathbf{b}^{(i)} \qquad (2)$$

Here $\mathbf{b}^{(i)} \in \{0, 1\}$ denotes whether the ith element is off or on; h_Z is the channel when all elements are off; and $\mathbf{h}^{(i)}$ is the effect of turning the ith element on. Here, we have folded the complexities of α_i into c_i to get $\mathbf{h}^{(i)}$. We in prove in §5.2 that having the ability to set *any* $|\mathbf{b}^{(i)}| \leq 1$ (i.e., fine-grained control over the phase) instead of only 0 or 1 gives only a factor-of-π advantage in optimizing $|h|$.

3.2 How Size Helps Communication

To get a qualitative understanding of the benefits of RFocus, we use some well known results from physics. First, RFocus can only control the portion of transmitted energy that falls on it. This is given by the solid angle, Ω, subtended by the surface on the transmitter. The Abbe diffraction limit tells us how well the surface can focus the energy at the receiver. If the surface modulates the incident energy perfectly, the spot onto which the energy is focused will have an area a proportional to $\lambda^2 \left(1 + 4\frac{d^2}{A}\right)$, where A is the area of the RFocus surface and d is the distance of the receiver from it [20]. Thus a fraction $\Omega/(4\pi)$ of the transmitted energy is now spread over an area $\approx a$, and hence peak energy will be proportional to Ω/a. RFocus works best when either the transmitter or the receiver is close to the surface, and therefore Ω is large (as a rule of thumb, about 3 meters for a surface 2-3 meters in length/width). Note, the Abbe diffraction limit is only an approximate result.

A/d^2 is proportional to the angle subtended by the surface on the receiver. There are two regimes depending on how large this angle is. If the radio is far away from the surface (i.e. $d^2 \gg A$), then $a \overset{\propto}{\sim} d^2/A$ and energy falls as $A\Omega/d^2$. This is still a $1/d^2$ fall, but the constant is improved by a factor of $A\Omega$. If
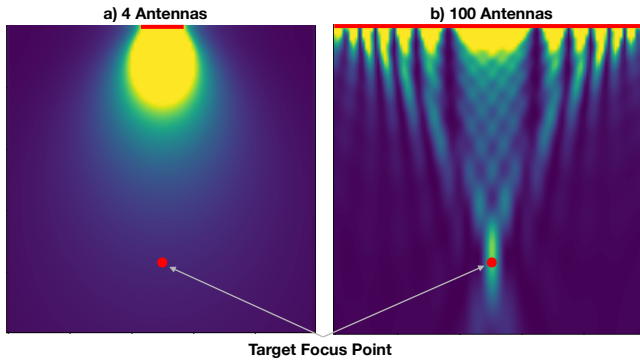
Figure 3: A simulation of how signal strength (brighter is higher) is distributed when an antenna array tries to maximize signal strength at the target point. The antennas are on the top horizontal line (red). With four antennas the signal quickly begins to diffuse. The 100 antennas subtend a large angle at the target, and are hence able to focus energy there, avoiding attenuation due to spreading.

the radio is closer, then $d^2 \sim A$, and the first term dominates. In this regime, a can be made quite small, on the order of a few λ^2. Hence almost all of the transmitted energy can be incident on the receiver.

In traditional beamforming, A is typically small, hence $d^2 \gg A$ and we are always in the first regime where the signal experiences a $1/d^2$ attenuation as it spreads out. The difference between the two regimes is illustrated in Figure 3.

4 Optimization Algorithm

The RFocus controller uses measurements from the radio endpoints to maximize signal strength at the receiver. In this section we first describe the challenges in measuring changes in the channel, and why we rely only on RSSI measurements (§4.1). Then we describe our algorithm (§4.2) before giving a preliminary theoretical analysis showing that it converges to a near-optimal solution (§4.3).

4.1 Measuring the Channel

A direct, but naive, method: When all the elements are turned off, the channel is h_Z, by definition. Ideally, to measure each $\mathbf{h}^{(i)}$, we could turn just the ith element on, and measure the difference from h_Z. But this change is usually too feeble to measure because an element is just a small piece of metal lying somewhere in the environment (§6.2.3). In many cases, we cannot statistically tell the difference between an element being on or off, even with hundreds of samples. This is also the reason why gradient descent to find an optimal configuration is untenable.

Boosting the signal. Each $\mathbf{h}^{(i)}$ may be small, but all the elements together can have a large effect. To make the change measurable, we generate several configurations of the surface by randomly choosing the state of each element. If we vary N elements, the change in the channel will have an expected magnitude of $O(\bar{h}\sqrt{N})$ (due to the central limit theorem), where $\bar{h}$ is the average magnitude of $\mathbf{h}^{(i)}$. This gives us a $O(\sqrt{N})$ boost in amplitude (and $O(N)$ boost in signal strength) over flipping just one element.

Challenges in measuring phase. One might consider a method that would measure the (complex) channel for many random configurations of the surface, and solve the linear equations to obtain all the $\mathbf{h}^{(i)}$. But this is difficult because it needs to measure changes in the *phase* of the channel, which commodity wireless devices don't provide. More importantly, the change in the channel, even after this $O(\sqrt{N})$ boost, remains small. Measuring small changes in phase is hard because the clocks of the transmitter and receiver are never perfectly synchronized, leading to carrier frequency offset (CFO) and jitter in phase measurements. Correcting for CFO over long periods of time is non-trivial, since CFO drifts over time.

Using RSSI. Our method uses only signal-strength measurements, ignoring phase information altogether. RSSI is not always an absolute metric, and may vary due to automatic-gain control changes or temperature changes at the amplifier. Hence we measure RSSI of a test state *relative* to a baseline state; e.g., to the all-zeros state where all elements are turned off. We call this quantity the *RSSI-ratio*.

4.2 RFocus's Optimization Algorithm

Algorithm 1 The majority voting algorithm (described below)

procedure MAJORITYVOTE($[\mathbf{b}_1,...,\mathbf{b}_K]$, RSSI)
    // RSSI[j] gives the measured RSSI-ratio for state $\mathbf{b}_j$
    $\mathbf{b}_\perp \leftarrow$ blank-vector   ▷ The final optimized state
    $m \leftarrow$ MEDIAN(RSSI)
    for $i := 0$ **to** N **do**
        VoteOn $\leftarrow 0$, VoteOff $\leftarrow 0$
        for $j := 0$ **to** K **do**
            if
                $(\mathbf{b}_j^{(i)} == 1$ **and** RSSI[j] $> m)$ **or**
                $(\mathbf{b}_j^{(i)} == 0$ **and** RSSI[j] $< m)$ **then**
                VoteOn $\leftarrow$ VoteOn $+ 1$
            else
                VoteOff $\leftarrow$ VoteOff $+ 1$
        $\mathbf{b}_\perp^{(i)} \leftarrow ($VoteOn $>$ VoteOff$)$
    return $\mathbf{b}_\perp$

Given the measured RSSI-ratio for a set of K random configurations, the idea is to compare the RSSI-ratio of each mea-

surement to the median value: if the RSSI-ratio when the ith element is on (off) is higher than the median, then we add a vote for the element to be on (off) in the optimized configuration. The ith element's optimized state is the state that received more votes for that element. Algorithm 1 gives the pseudo-code for how the controller determines the optimized state for each bit, i.

We use the median instead of the mean of the RSSI-ratio samples because it is more robust to one-shot noise, which may occur due to people moving or amplifier changes. We expect this algorithm to be robust to element failure and occluded elements, since elements' states are determined independently of each other. We evaluate how the received signal strength of the optimized configuration degrades with element failures in Section 6.4.

4.3 Theoretical Analysis

We show that if each $\mathbf{h}^{(i)}$ is small and the collection has uniformly-distributed phases, the optimization algorithm will find a near-optimal solution with high probability.

4.3.1 Assumptions

The proof relies on the following assumptions:

Assumption 1. *The phases of $\mathbf{h}^{(i)}$ are uniformly distributed random variables in $(-\pi,\pi]$.*

This is reasonable because the paths to different elements are of different lengths, so all phases are equally likely to be represented. To remove any spatial correlations, we can randomize the indices of $\mathbf{h}^{(i)}$.

Assumption 2. *For uniformly random $\mathbf{b}$, $|\mathbf{b} \cdot \mathbf{h}| \ll h_Z$ with high probability.*

The average of $\mathbf{h}^{(i)}$ is 0 when phases are uniformly random. Hence, assumption 1 implies that for uniformly random $\mathbf{b}$, $|\mathbf{h} \cdot \mathbf{b}|$ is $O(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}|\mathbf{h}^{(i)}|)$ due to the central limit theorem.

Though $|\mathbf{h} \cdot \mathbf{b}|$ is small for random $\mathbf{b}$, an *optimal* assignment to $\mathbf{b}$ can cause $|\mathbf{h} \cdot \mathbf{b}|$ be large relative to $|h_Z|$, since it will be $O(N)$ times bigger than the average $|\mathbf{h}^{(i)}|$ (see theorem 2). On the other hand, for random $|\mathbf{b}|$, $|\mathbf{h} \cdot \mathbf{b}|$ is only $O(\sqrt{N})$ times bigger than the average $|\mathbf{h}^{(i)}|$.

Assumption 3. *$|\mathbf{h}^{(i)}|$ is bounded above by a constant, even as $N \to \infty$.*

With this assumption, $\left|\mathbf{h}^{(i)}\right| \ll \sum_{j=i}^{N}\left|\mathbf{h}^{(j)}\right|$, capturing the intuition that, since all elements are small, there is no dominating element. Many elements must contribute to create a large effect.

Note that an RFocus element is only half a wavelength long when on. This implies that, individually, an element cannot be very directional [9]. Hence, the reflections are not specular. Rather, each element diffracts waves to a wide range of angles. Hence, we would expect all the $\mathbf{h}^{(i)}$ to have a similar (likely,
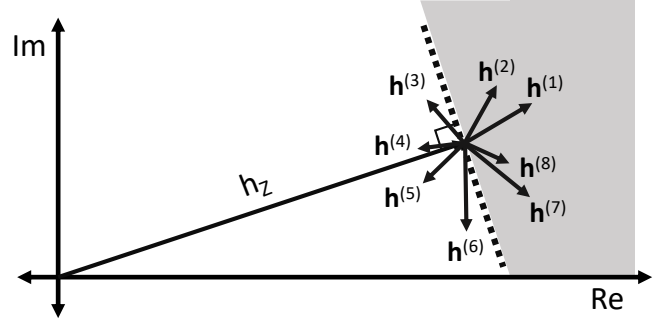

Figure 4: h_Z and the $\mathbf{h}^{(i)}$ in the complex plane. To maximize signal strength, we should pick the $\mathbf{h}^{(i)}$ in the shaded region and remove the others.

small) magnitude. However, acting together, the elements form a large structure which which c which and can beconfigured to achieve directionality (§3.2). In addition, we assume that equation 2 is accurate, which we later verify experimentally (see §6.2.1). Our analysis is in the limit where the number of elements and the number of measurements go to infinity.

4.3.2 Proof Outline

Consider h_Z and $\mathbf{h}^{(i)}$ on the complex plane as shown in Figure 4. The signal strength is $|h_Z + \sum_i \mathbf{h}^{(i)}\mathbf{b}^{(i)}|^2$. If all the elements are turned on, the $\mathbf{h}^{(i)}$ interfere destructively with each other. To increase signal strength, we should eliminate this destructive interference. To do so, we should pick a half-plane (the dotted line), turn on all elements on one side of it and turn the others off. We show in Lemma 1 that the optimal solution takes this form.

Next, if the phases of $\mathbf{h}^{(i)}$ are randomly distributed, it doesn't matter which half-plane we pick; they will all produce roughly the same $|\mathbf{b} \cdot \mathbf{h}|$ (Lemma 2). Hence, we can maximize alignment with h_Z by picking the half-plane perpendicular to it (pictured in the figure).

The controller needs to determine for each element i if $\mathbf{h}^{(i)}$ is aligned with h_Z. Or equivalently, whether $\Re(\mathbf{h}^{(i)}h_Z^*) > 0$ where $\Re(\cdot)$ denotes the real component of a complex number and $*$ denotes complex conjugation. The algorithm takes advantage of the fact that if $\Re(\mathbf{h}^{(i)}h_Z^*) > 0$, the signal strength is slightly more likely to be higher than the median signal strength when $\mathbf{h}^{(i)}$ is turned on (Theorem 1).

4.3.3 A Near-Optimal Solution

We show that all the chosen $\mathbf{h}^{(i)}$ in an optimal solution lie on one side of a line in the complex plane, formalized as follows:

Lemma 1. *Under assumptions 2 and 3, let $\mathbf{b}_{OPT}$ be an optimal state assignment. Then, $\mathbf{b}_{OPT}^{(i)} = 1$ if and only if $\Re(\mathbf{h}^{(i)} \cdot H(\mathbf{b}_{OPT})^*) > 0$, where $H(\mathbf{b}) = h_Z + \mathbf{h} \cdot \mathbf{b}$.*

Proof. If this were not the case, we could flip $\mathbf{b}_{OPT}^{(i)}$ to get

$$|H(\mathbf{b}_{OPT}) - \mathbf{b}_{OPT}^{(i)}\mathbf{h}^{(i)} + \left(1-\mathbf{b}_{OPT}^{(i)}\right)\mathbf{h}^{(i)}|$$
$$= \left|H(\mathbf{b}_{OPT}) + \left(1-2\mathbf{b}_{OPT}^{(i)}\right)\mathbf{h}^{(i)}\right| \quad (3)$$
$$\geq |H(\mathbf{b}_{OPT})|$$

To prove the last inequality, note

$$\left|H(\mathbf{b}_{OPT}) + \left(1-2\mathbf{b}^{(i)}\right)\mathbf{h}^{(i)}\right|^2 - |H(\mathbf{b}_{OPT})|^2$$
$$= \left|\left(1-2\mathbf{b}^{(i)}\right)\mathbf{h}^{(i)}\right|^2 + 2\Re\left(\left(1-2\mathbf{b}^{(i)}\right)\mathbf{h}^{(i)} \cdot H(\mathbf{b}_{OPT})^*\right) \quad (4)$$
$$\approx 2\Re((1-2\mathbf{b}^{(i)})\mathbf{h}^{(i)} \cdot H(\mathbf{b}_{OPT})^*)$$
$$\geq 0$$

The second-last approximation follows because $|\mathbf{h}^{(i)}|$ is bounded and hence $|\mathbf{h}^{(i)}| \ll h_Z \leq |H(\mathbf{b}_{OPT})|$. The last inequality holds if our condition on $\mathbf{b}_{OPT}^{(i)}$ is not satisfied. If so, the solution can be improved, which contradicts our assumption that $\mathbf{b}_{OPT}$ is optimal. $\square$

The next lemma states that the solution pictured in Figure 4 is close to optimal with high probability as the number of elements $N \to \infty$.

Lemma 2. *Let $\mathbf{b}_{OPT}$ be the optimal assignment that maximizes $|h_Z + \mathbf{h} \cdot \mathbf{b}|$ and $\mathbf{b}_\perp$ be such that the ith component $\mathbf{b}_\perp^{(i)} = 1$ if and only if $\Re(\mathbf{h}^{(i)} \cdot h_Z^*) > 0$. As $N \to \infty$, if assumptions 1 and 3 hold, then $|H(\mathbf{b}_{OPT})|/|H(\mathbf{b}_\perp)| < 1+\varepsilon \ \forall \varepsilon > 0$, with high probability.*

The proof is in Appendix A. We give the outline here. Let $\mathbf{b}_\theta$ be the state such that $\mathbf{b}_\theta^{(i)} = 1$ iff $\Re(\mathbf{h}^{(i)}e^{-j\theta}) > 0$. We show that, since the phases of $\mathbf{h}$ are uniformly distributed, it does not matter what θ we pick. That is, $\max_\theta |\mathbf{b}_\theta \cdot \mathbf{h}|$ isn't very different from $\min_\theta |\mathbf{b}_\theta \cdot \mathbf{h}|$. Hence, we can pick θ to be the phase of h_Z to prove the theorem.

4.3.4 Analysis of the Algorithm

We prove that Algorithm 1 finds the near-optimal solution discussed in Lemma 2.

Theorem 1. *Let assumptions 1, 2, and 3 hold. Then, as $N \to \infty$ and $K \to \infty$, the configuration returned by Algorithm 1 finds a near-optimal solution.*

Proof. According to Equation 2, when we randomly vary $\mathbf{b}^{(i)}$, h becomes a random variable, H, with mean $h_Z + S/2$, where $S = \sum_{i=1}^N \mathbf{h}^{(i)}$ is the sum of the contributions of the rest of the elements. The $S/2$ term appears because we include each element with probability $1/2$. Figure 5(a) shows this probability distribution. Consider an element i that hasn't yet been fixed. If we condition the probability distribution on the ith element being on, then the PDF shifts by $\mathbf{h}^{(i)}/2$ as
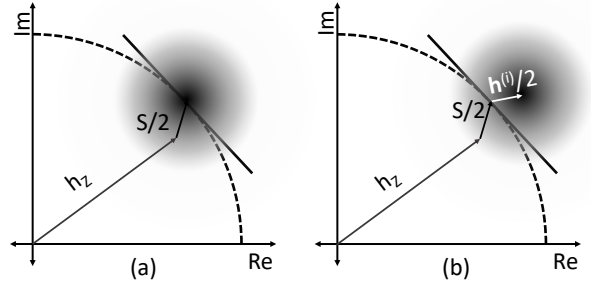

Figure 5: (a) Shows the probability density function of the channel when all bits are chosen uniformly at random. (b) Shows the PDF conditioned on the ith bit being 'on'. The dashed circle is centered at the origin with radius $|h_Z + S/2|$, where $S = \sum_{i=1}^N \mathbf{h}^{(i)}$. Depending on which side of the circle h_i takes the mean, the mean magnitude will be greater or lesser than that of the unconditional PDF.

shown in Figure 5(b), because the i^{th}element's value is fixed in these samples (it shifts by $\mathbf{h}^{(i)}/2$ and not $\mathbf{h}^{(i)}$ since we have already included the other $\mathbf{h}^{(i)}/2$ in $S/2$). If we condition on the element being off, then the mean shifts by $-\mathbf{h}^{(i)}/2$.

The central limit theorem implies that $H/\sqrt{N}$ is Gaussian as $N \to \infty$. Hence, if $\mathbf{h}^{(i)}/2$ (or $-\mathbf{h}^{(i)}/2$) shifts the mean to be outside the circle, then when the element is on (or off) the RSSI-ratio is more likely to be greater than the unconditional mean. If $\mathbf{h}^{(i)}/2$ shifts the mean inside the circle, then the opposite holds. Here, we use assumption 2, which implies that H isn't shifted far from $h_Z + S/2$. Hence, as $K \to \infty$, Algorithm 1 determines with confidence tending to 100%, whether the conditional mean is inside or outside the circle, by looking at RSSI alone. Note that mean equals median in this case.

Assumption 2 implies that $|\mathbf{h}^{(i)}| \ll |h_Z|$. Thus, the mean shifts outside the circle if and only if it shifts to the outer side of the tangent line shown in figure 5. That is, $\mathbf{b}_\perp^{(i)} = 1$ if and only if $\Re\left(\mathbf{h}^{(i)} \cdot (h_Z + S/2)^*\right)$. But $h_Z + S/2 \approx h_Z$, because of assumption 2. Hence $\mathbf{b}_\perp^{(i)} = 1$ if and only if $\Re\left(\mathbf{h}^{(i)} \cdot h_Z^*\right) > 0$. From Lemma 2, we know that this produces a near-optimal solution. $\square$

To a first order approximation, the RSSI-ratio is linear in $\mathbf{b}$.[2] Hence linear regression is an alternative to majority voting. Majority voting has three advantages. First, unlike linear regression, majority voting is robust to outliers. Second, it is conceptually and computationally simpler, which is particularly important when the controller is an embedded device. Third, when the votes for a particular element determine its value with high confidence (say, $>95\%$), it can be fixed for the rest of the training

---

[2]RSSI-ratio, $|h/h_Z| = |1 + \mathbf{b} \cdot \mathbf{h}/h_Z| \approx 1 + \sum_i \Re(\mathbf{b}^{(i)}\mathbf{h}^{(i)}/h_Z)$ because $|\mathbf{b} \cdot \mathbf{h}| \ll h_Z$ for random $\mathbf{b}$ and $|1+x| \approx 1 + \Re(x)$ for small x.

period. This allows the channel to be improved even before the training ends. Fixing elements also removes the confounding effect they have on determining values for other elements. An earlier version of this paper used this idea [3]. However due to a performance optimization (§6.1), we do not implement it here.

5 Antenna Array Design

We have made two key design choices in the physical design of our antenna array. First, each element has only two states: one that reflects the signal, another that lets it through. Second, each element is half a wavelength tall and $1/10$ of a wavelength wide. In this section we explore the tradeoffs in these choices.

5.1 How Big Should Each Element Be?

An array with many small antennas gives better control over the reflected signal, while one with fewer but larger antennas is cheaper and simpler. If the inter-antenna spacing is larger than half a wavelength, there will be grating lobes where the signal is sent in directions other than the one desired [4].

Designing antennas that are much smaller than a wavelength[3] is challenging. Small antennas are either inefficient, absorbing only a small fraction of incident energy, or they are efficient only over a small bandwidth [21, 31]. Further, when placed close to each other, antennas interact strongly with each other in a way that is often hard to model.

Fortunately, a well understood result states that controlling the reflected wave at a granularity finer than half a wavelength gives only marginal benefits. Consider two infinite parallel planes a distance d apart, separated by a homogeneous medium. Variations in electric/magnetic fields in one plane that are faster than once per wavelength, will have a negligible effect on the fields on the other plane; it decays exponentially with d [24]. Hence any fine-grained variations we introduce in the surface will be lost as soon as the signal propagates a few wavelengths in either direction. Thus we can design an array with antennas comparable to a wavelength and still get most of the benefits. Metasurfaces [6] may enable a more efficient design, but that is out of scope of this work.

5.2 How Many States for an Element?

We chose elements that can be in only one of two states. But we could have chosen a design that offers greater control. Ideally we would be able to control the exact phase and amplitude with which each element reflects its energy. In terms of our model in equation 2, we would have been able to set any $\mathbf{b}^{(i)} \in \mathbb{C}, |\mathbf{b}^{(i)}| \leq 1$, instead of being restricted to $\mathbf{b}^{(i)} \in \{0,1\}$. Denote the amount of energy that can be directed by the surface in the two cases as h_{IDEAL} and h_{REAL}. In the ideal

---

[3]These are called "electrically small antennas".

---

system, we would be able to align the phases of all $\mathbf{h}^{(i)}$ to get $|h_{IDEAL}| = \sum_i |\mathbf{h}^{(i)}|$. We show here that $|h_{REAL}| \geq |h_{IDEAL}|/\pi$.

Here $|h_{REAL}|$ and $|h_{IDEAL}|$ include only the signal due to the surface ($|\mathbf{h} \cdot \mathbf{b}|$) and not the rest of the environment (h_Z). To maximize signal strength, we would need to align the phases with h_Z too. If we assume that phases are random (assumption 1), we can also prove the result when $|h_{REAL}|$ and $|h_{IDEAL}|$ include h_Z. We discuss this as a corollary of lemma 3 in Appendix A.

Theorem 2. *Under assumptions 2 and 3, $|h_{REAL}| \geq \frac{|h_{IDEAL}|}{\pi}$ as $N \to \infty$.*

Proof. Define $A = \int_{-\pi}^{\pi} \sum_{i=1}^{N} |\Re(\mathbf{h}^{(i)} \cdot e^{-j\theta})| d\theta$. $\Re(\cdot)$ denotes the real part of a complex number. The variable A expresses the sum of components of $\mathbf{h}^{(i)}$ along angle θ and integrates over all θ. Each θ corresponds to a perpendicular to a half-plane, as discussed before in Lemma 1. At least one of these, θ_0, corresponds to the optimal half-plane, wherein the optimal solution contains all the $\mathbf{h}^{(i)}$ such that $\Re(\mathbf{h}^{(i)} \cdot e^{-j\theta_0}) > 0$. These contribute $\mathbf{h}^{(i)}$ toward h_{REAL}. Thus,

$$|h_{REAL}| \geq \sum_{i=1}^{N} \max\left\{0, \Re\left(\mathbf{h}^{(i)} \cdot e^{-j\theta_0}\right)\right\} \geq \frac{1}{2} \sum_{i=1}^{N} \left|\Re\left(\mathbf{h}^{(i)} \cdot e^{-j\theta_0}\right)\right| \tag{5}$$

The first inequality holds because the absolute value (LHS) is greater than the real component (RHS). The second inequality holds because otherwise we could have chosen $\pi + \theta_0$ and obtained a better $|h_{REAL}|$. Hence $|h_{REAL}| \geq \frac{1}{2} \max_{\theta \in [\pi, \pi)} \sum_{i=1}^{N} |\Re(\mathbf{h}^{(i)} \cdot e^{-j\theta})| \geq \frac{1}{2} \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{i=1}^{N} |\Re(\mathbf{h}^{(i)} \cdot e^{-j\theta})| d\theta = \frac{A}{4\pi}$, since the maximum is greater than the average.

Separately, we can rearrange the sum as $A = \sum_{i=1}^{N} \int_{-\pi}^{\pi} |\Re(\mathbf{h}^{(i)} \cdot e^{-j\theta})| d\theta = \sum_{i=1}^{N} |\mathbf{h}^{(i)}| \left(\int_{-\pi}^{\pi} |\cos\theta| d\theta\right)$. The second step is possible, because $\cos\theta$ expresses the dot product of $\mathbf{h}^{(i)}$ over a unit complex number with phase θ. Since we are integrating over all angles, it doesn't matter which angle we start from. Now we can evaluate the integral to get $A = 4 \sum_{i=1}^{N} |\mathbf{h}^{(i)}| = 4|h_{IDEAL}|$. Since $|h_{REAL}| \geq \frac{A}{4\pi}$, $|h_{REAL}| \geq \frac{|h_{IDEAL}|}{\pi}$. □

Alternate designs. We could design an alternate system where $\mathbf{b}^{(i)} \in \{-1, 1\}$ instead of $\mathbf{b}^{(i)} \in \{0, 1\}$. That is, the elements either reflect signal directly or with a 180° phase difference. In this case, we'd get a $2/\pi$-approximation, which is better than the $1/\pi$ factor we get now. This is because we would have had $\max\{\Re(\mathbf{h}^{(i)} \cdot e^{j\theta_0}), \Re(\mathbf{h}^{(i)} \cdot e^{-j\theta_0})\}$, which removes the $1/2$ factor in equation 5. On the other hand, while our setup can function as both a mirror and a lens (i.e. it can be selectively transparent), this alternate design can only operate as one of the two. A four-state system could potentially offer the benefits of both, a good topic for future work.

5.3 Our Design

Our surface design consists of a panel of metal rectangles of size $\lambda/4 \times \lambda/10$ as shown in Figure 6, where λ denotes
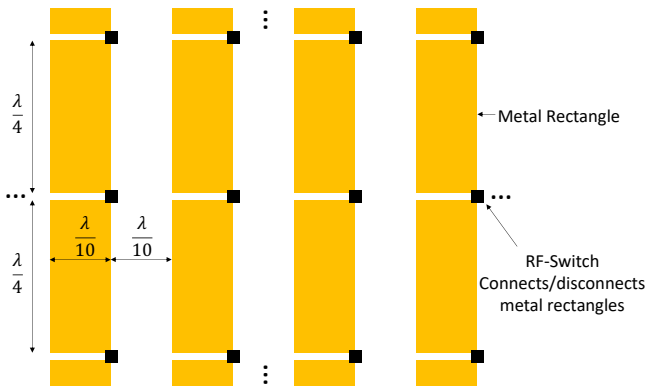

Figure 6: Schematic of the design of our antenna array. This array of rectangles continues in both directions.

the wavelength. They are connected by RF switches, which determine whether or not the rectangles are connected (the switches are placed off-center for practical PCB-design reasons). This design works only for vertically polarized radiation. It can be generalized to all polarizations by having an identical copy perpendicular to this one.

There are two principles of operation. First, the rectangles by themselves are small and hence interact weakly with radiation. If the switches are all off, the surface would be semi-transparent. However, when an RF switch is turned on, it joins two rectangles to form a $\lambda/2 \times \lambda/10$ rectangle. This forms a half-dipole antenna and interacts strongly with incident radiation. We made the strips wide to support a wider bandwidth of operation, since the width allows more modes of oscillation.

Second, if a plate of metal has small holes in it, then radio behaves as if the holes weren't there. This effect is used in Faraday cages, such as those in microwave ovens. A common rule-of-thumb says that the holes need to be smaller than $\lambda/10$, the size of the gaps between rectangles. When switches in adjacent columns are turned on, their rectangles will behave as a continuous sheet of metal, rather than individual columns, almost completely reflecting incident waves. Because neighbors act in a simple way, we expect that the neighbors' state wouldn't change the phase of the currents induced in the rectangle, only the magnitude. Hence the linear model in Equation 2 is approximately correct.

The above reasoning is merely the conjecture that motivated our design. We conduct two experiments to partially validate it. §6.2.1 demonstrates that the linear model is approximately correct, and §6.2.2 shows that the surface can significantly change its opacity. Validating this design in an anechoic chamber would offer more insights. We leave that for future work.

6 Evaluation

6.1 Experiment Setup

Our antennas are fabricated on custom printed circuit boards, with 40 antennas per board. We mount 80 of these boards on two frames and place it next to a wall in our lab. Since metal would interfere with the antennas, the frames are partially built from acrylic. The boards are connected in series with a single SPI serial-to-parallel bus composed of shift registers, allowing our Raspberry Pi controller to serially set the state of each individual element.

To set element state, the controller pushes values through the shift registers at 1 Mbit/s. Higher speeds are impeded by distortion due to the long wires in our setup. Hence pushing a new random state to the surface is the most time-consuming part of measuring it. To alleviate this problem, instead of pushing 3200 bits for each state, we push just 8 bits at a time, which shifts all the bits in the surface by 8 positions. Though the new state isn't completely fresh, each element still gets a new random value that is independent of the previous one. If we were to solve for $\mathbf{h}^{(i)}$ from the system of linear equations, the measurement matrix is still full rank. Hence we expect Algorithm 1 to work. This modification reduces the mean time to measure a state from 7 ms to 0.22 ms. Better engineering will improve performance further. To measure the channel, we use a USRP programmable radio operating in the 2.4 GHz ISM band.

Measurement method. The results in the main evaluation (Figures 10 and 11) are reported as the ratio of the signal strength when RFocus is in its optimized state to the strength when the surface is not present (i.e., *it is physically removed*, not just set to off). We measure this ratio as follows. Let R_I be the improved signal strength, R_0 be the signal strength when the surface is in the all-zeros state (i.e. it is semi-transparent), and R_B be the baseline signal strength when RFocus isn't present. We separately measure R_I/R_0 and R_0/R_B and multiply them to obtain R_I/R_B as desired.

After optimizing the state, we measure R_I/R_0 100 times by repeatedly flipping between the two states every 5 ms and calculating the mean improvement. We find that 1σ error is $< 2\%$ in all cases. To measure R_0/R_B, we first measure R_0. Then we remove the surface and measure R_B. We do this over ≈ 10 transmitter positions in a neighborhood around the original point, and use the median R_0/R_B to correct R_I/R_0. The median helps account for the *systematic* effects of the RFocus surface. It ignores random fading effects that may occur since the RFocus surface, people, or objects will have moved in the meanwhile. The mean absolute correction we had to introduce was 0.6 dB, while the maximum was 1.5 dB.

Total prediction error	Error due to noise
5.4%	2.0%

Figure 7: A linear model predicts the channel due to a state with 5.4% accuracy. If the surface were perfectly linear, the error would have been 2.0% due to noise. Hence the RFocus is approximately, but not fully, linear.

6.2 Microbenchmarks

6.2.1 Linearity

Our analysis of the optimization algorithm (see §4.3) assumes that the elements are not coupled with each other. Hence we used a linear model formalized in Equation (2), where $\mathbf{h}^{(i)}$ are independent of the states of the other elements. We wouldn't expect elements that are far away from each other (say, more than a wavelength) to couple strongly, especially since the elements are flat and radiate perpendicularly away from the surface. However, we packed the elements tightly to more densely cover the surface and beamform more precisely. We hypothesized that the non-linearity due to any resulting coupling will be small (§5.3). We now test this hypothesis.

We prepare several random "test-triples" of states of the form $(\mathbf{b}_A, \mathbf{b}_B, \mathbf{b}_{AB})$. Let (h_A, h_B, h_{AB}) be the corresponding channels. The states are drawn uniformly from the triples where $\mathbf{b}_A \& \mathbf{b}_B = 0$ and $\mathbf{b}_{AB} = \mathbf{b}_A | \mathbf{b}_B$; here states are represented as bit-strings and $\&$ and $|$ are bitwise operators. That is, no element is on in both $\mathbf{b}_A$ and $\mathbf{b}_B$, and any element that is on in either $\mathbf{b}_A$ or $\mathbf{b}_B$ is also on in $\mathbf{b}_{AB}$.

If the linear model in Equation (2) is correct, then $h_A/h_Z + h_B/h_Z - 1 = h_{AB}/h_Z$. If an element's effect on the channel ($\mathbf{h}^{(i)}$) depends on whether its neighbors are on, this relation will not hold, since many more pairs of neighbors are on in $\mathbf{b}_{AB}$ than in $\mathbf{b}_A$ or $\mathbf{b}_B$. Since the neighborhoods are substantially different, $\mathbf{h}^{(i)}$ will be different for different states if our assumption is wrong.

Hence, to test linearity, we measure $(|h_A/h_Z|, |h_B/h_Z|, |h_{AB}/h_Z|)$ for several random states and test our ability to predict $|h_{AB}/h_Z|$ given $|h_A/h_Z|$ and $|h_B/h_Z|$ using the above relation. Using $|\cdot|$ doesn't invalidate the relation since for any random state X, $h_X/h_Z = |1 + \Delta h_X/h_Z| \approx 1 + \Re(h_X/h_Z)$ since $\Delta h_X \ll h_Z$. We measure each state 100 times. If Equation (2) were correct, the error in prediction due to measurement noise would have been 2.0%. Our error is 5.4% as shown in Figure 7. We conclude that although RFocus is not perfectly linear, the nonlinearities are small. They *are*, however, large enough that linear regression on $|h_X/h_Z|$ doesn't produce a good predictor.

6.2.2 Controllability and Bandwidth

We expect that the RFocus surface can control its opacity to radio. To test this, we kept the surface in between two wide-bandwidth directional (Vivaldi) antennas pointed at each other.
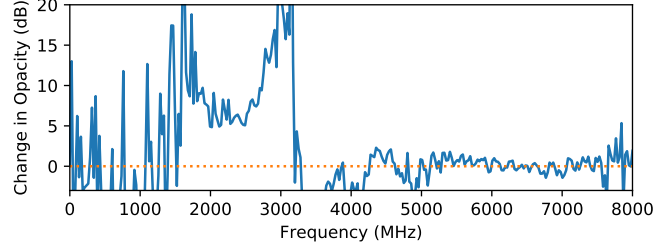

Figure 8: The ability to control the surface's opacity to radio as a function of frequency. This also indicates RFocus's bandwidth of operation.
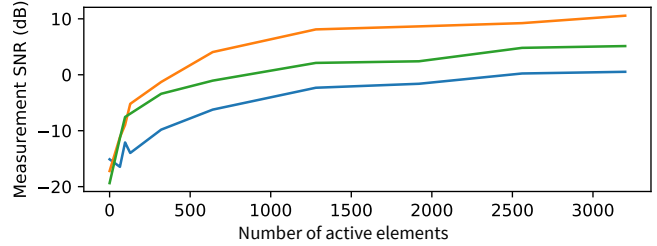

Figure 9: Measurability of the effect random configurations of the RFocus surface on the channel, as a random subset of them are deactivated, for three representative transmitter locations. It demonstrates why boosting (§4.1) is important for RFocus to function.

Using a Vector Network Analyzer (VNA), we compare the signal strength between the antennas when all the elements are turned on to when they are all turned off. We expect that, when the elements are all on, the surface will be much more opaque to radiation, reflecting a large fraction of it. The ratio of signal strength in these configurations is shown in Figure 8: it is consistently greater than 6 dB between 1600 and 3100 MHz. Hence, RFocus can change its opacity by over 75% over a large bandwidth. The peak is closer to 3000 MHz, where the change is well over 10 dB (90% control). But all of our other results are in the 2450 MHz ISM band, in order to conform to FCC rules. Frequency of operation can be tuned by scaling the sizes of the components. (Antenna design is not the focus of this paper.)

The *y*-axis is cropped at -3 dB for clarity in showing our frequency range of interest. The change falls after 3000 MHz because our RF switch is only rated up to that frequency. At < 1500 MHz, the rectangles, even after joining, are too small to interact with radiation.

6.2.3 Measurability

To find a good configuration, the controller needs to measure $\mathbf{h}$. However, the effect of each individual element, $\mathbf{h}^{(i)}$, is tiny and hard to measure. To aid measurement, we vary all elements randomly and at once (§4.1). This gives us an $O(\sqrt{N})$ boost in
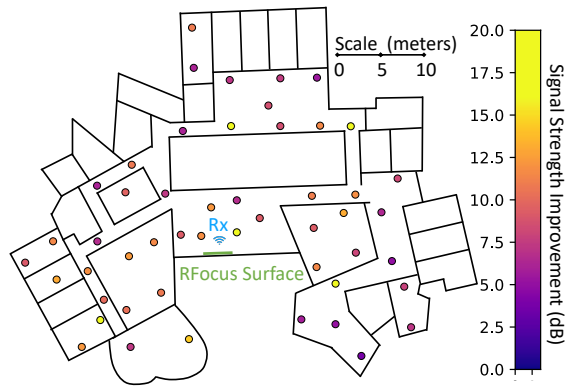
Figure 10: Ratio (in dB) of signal strength without and without the surface. The colored circles represent the improvement when the transmitter was placed at those locations, with the receiver and RFocus placed as shown. The CDF of improvements are shown in Figure 11.

Figure 11: CDF of Improvement in the signal strength and channel capacity

the change in the channel amplitude ($O(N)$ in signal strength), which is easier to measure. To experimentally study the effect of this boost, we compute the Signal to Noise Ratio (SNR) of the *change in the channel due to a random state* as a function of the number of elements in the array. We artificially reduce the size of our array, by deactivating a random subset of it.

We choose 165 different random configurations, and measure the RSSI-ratio for each configuration ≈ 1000 times. The definition of "measurability SNR" is as follows. The "signal" in SNR is the variance in the average RSSI-ratio across all configurations, and the "noise" is the average variance in the RSSI-ratio measurements within each individual configuration. With the receiver and RFocus surface in the locations shown in Figure 10, we place the transmitter at three representative locations covering the full range of signal strengths.

Figure 9 shows the measurability SNR as a function of the number of active elements for these points. We can see that SNR is increases when more elements are varied. Note that when only a few elements are varied, the SNR is very low. This is why boosting the signal by varying all elements at once is critical, especially when the transmitter is far from the receiver.

The impact of random configurations on the channel is still small, because elements can interfere destructively. An optimized state eliminates destructive interference and produces an $O(N)$ effect on the channel ($O(N^2)$ in signal strength). Hence it produces significant gains in signal strength, even though a random state has little impact.

6.3 Signal Strength Optimization

We placed the receiver and the RFocus surface at a constant location as marked in Figure 10. Then we placed the transmitter at various positions in an indoor environment (our lab), and ran
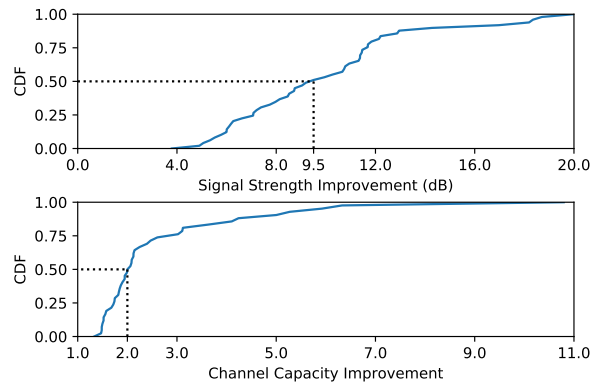
the optimization algorithm to maximize signal strength at the receiver. We measure the ratio of the improved signal strength to the signal strength when the surface is physically removed. We plot these on a map in Figure 10. The corresponding CDFs are shown in Figure 11. RFocus provides benefits throughout an entire floor of our building. The minimum, median and maximum improvements across all locations are 3.8 dB, 9.5 dB, and 20.0 dB, respectively.

The five points with the highest improvement had among the smallest unimproved RSSI. Further, they appear in apparently random locations, which makes us postulate that their unimproved RSSI is low due to fading, where destructive interference drastically decreases signal strength. Due to its size, RFocus is unlikely to suffer from fading at all its antennas. Hence, RFocus provides significant benefits for weak channels in such situations.

RFocus's large area allows it to focus energy from the transmitter to the receiver. This is particularly helpful when the transmitters are power constrained, since even a "whisper" will be "heard" clearly at the receiver. Yet, interference does not increase because the transmit power isn't increased. This could enable a new regime of low-power, high throughput IoT sensor devices. Whenever a sensor wants to transmit, it can ask the controller to tailor the surface for that particular endpoint (using a brief high-power transmission). Then it can make its high-bandwidth transmission at low power. Since sensors tend to be stationary, the same trained configuration can be used for extended periods of time. We evaluate this scenario in 6.4.1. Once trained, setting a state takes ≈ 3 ms.

Alternately, radios can use the additional signal strength to increase their channel capacity to obtain higher bitrates, as shown in Figure 11. Channel capacity increases logarithmically with signal strength, and hence the improvement decreases with increasing transmitting power. We set our transmitter to the highest power level (≈ 20 dBm) and adjust the receiver's gain accordingly to avoid saturation.
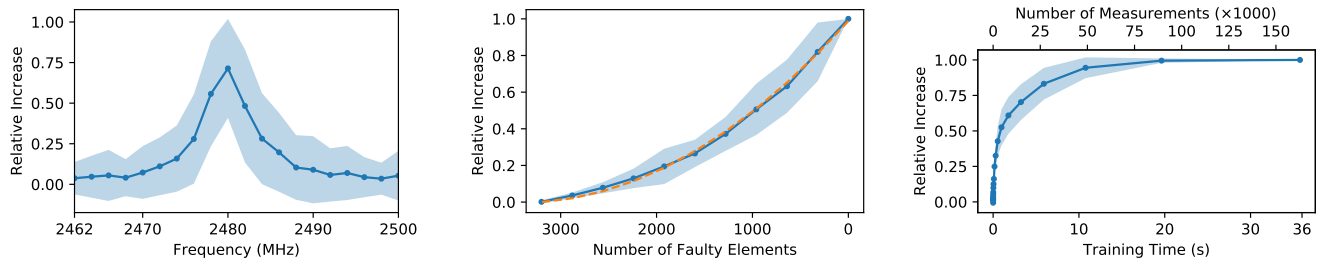
(a) Improvement for different frequencies, when RFocus was optimized for 2480 MHz.

(b) Improvement when some antennas are faulty and flip randomly. Dotted line shows the best quadratic fit.

(c) Improvement as a function of training time (and number of measurements)

Figure 12: To understand the trends as we vary different parameters, we plot the mean $\pm$ standard deviation of the improvement across all locations pictured in Figure 10. To normalize for differences in absolute improvement for different locations, we plot $\frac{I-1}{Max(I)-1}$. I is the ratio of the improved signal strength to the baseline strength. We normalize relative to the maximum improvement (across all values of the parameter) for that location. Note: $I = 1$ means no improvement. Hence we subtract 1 in our metric, so that positive values denote an improved channel.

6.4 Understanding the Improvement

Neighboring frequencies. Our optimization algorithm only seeks to improve the signal strength at a single frequency (we leave generalizing to multiple frequencies as future work). Nevertheless, we find that it also provides improvement in a 10 MHz neighborhood as shown in Figure 12(a). Multi-path rich environments often have very different signal strength for different frequencies, due to fading [30]. To maximize gain, RFocus could preferentially improve frequencies with the lowest signal strength.

In some cases the frequency that improved the most is different from the target frequency. This could be because that frequency had a much lower baseline signal strength in the un-improved channel due to fading. Hence, the center frequency doesn't have a value of 1. Far away from the target frequency, sometimes the signal strength decreases slightly due to the surface.

Fault tolerance. RFocus is intended to be a large array of inexpensive elements. Hence individual elements are expected to fail, and the system must be robust to failure. To test this, we artificially make a random subset of elements flip randomly, and not as per the controller's instructions, during both training and testing. Figure 12(b) shows the improvement against the number of 'faulty' elements.

Our model of the system suggests that the signal strength increases quadratically with the number of elements (§3.2). This is because each element contributes linearly to the channel amplitude, and the signal strength is the square of the amplitude. Under this model, we expect the normalized mean plotted to have the form $\alpha n^2 + \beta n$, where n is the number of non-faulty elements, and (α, β) are constants. This is the best-fit line pictured.

Note that, this does not imply that gain increases quadratically with the area of the surface, since a larger surface will have more elements that are far away from the source, and hence a smaller $\mathbf{h}^{(i)}$. In the above experiment, we don't change the distribution of $\mathbf{h}^{(i)}$ since we disable a random subset of elements. Hence, the quadratic model works.

Optimization speed. To understand the rate at which the optimization progresses, we plot the signal strength improvement as a function of the training time/number of measurements. As shown in Figure 12(c), 50% of the improvement occurs within 1 sec, with 4500 measurements and 90% improvement occurs within 10 seconds. Note, RFocus has 3200 elements, and we'd expect to need at-least 3200 measurements before the problem is well determined, even ignoring noise. RFocus exploits additional measurements to contend with the fact that measurements are noisy (§6.2.3). RFocus provides some (smaller) benefit with fewer than 3200 measurements, since, at any point in time, RFocus has a hypothesis state that it believes is optimal.

6.4.1 Stability Across Time

Once optimized, how long does an optimized state work? How often do we need to re-optimize our state? To test this, we keep our transmitter in our office on a typical workday, where around 10 people work. We optimized the state at 10AM and measured the improvement due to *that* state till 4PM that day. The top graph in figure 13 shows this. The initial improvement, shown as a dotted line, was 12.6 dB for about 3 minutes before it increased (without re-optimizing the state) for the next ≈ 20 minutes, and then decreased again. We also do a 100 second experiment (in a different location), where two people walk between the endpoints and the surface, intentionally trying to block the direct paths. The bottom graph in figure 13 shows this. Note, the people move only in the middle portion. At other times, the improvement reverts to the baseline.

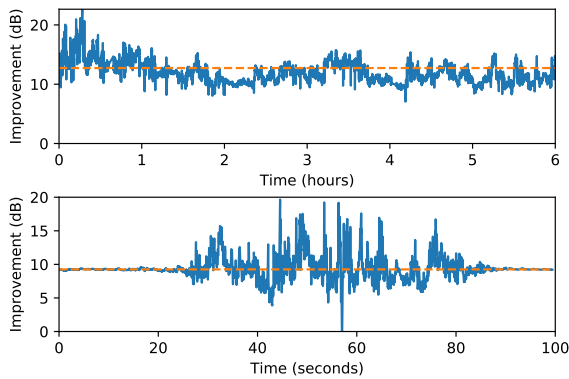RFocus's optimized state is robust to motion in the

Figure 13: Improvement across time due to a state optimized in the beginning of the time window. This shows that the same optimized state can be used for a long time. The dotted line shows the improvement immediately after the state was optimized. Note the unit change in the x-axis (hours/seconds).
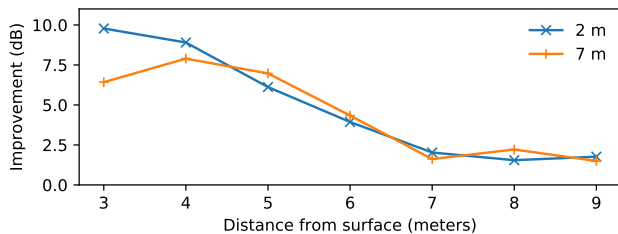

Figure 14: Signal strength improvement when *both* transmitter and receiver are at increasing distances from the surface. In other experiments, the receiver stays close to the surface. The radios endpoints are 2 m and 7m away from each other, in front of the surface, on the same side.

environment. Since optimization takes only 1 to 10 seconds, endpoints can benefit from RFocus for a large fraction of time. RFocus does not support continuous motion and needs to re-optimize the state if the endpoints move.

6.4.2 Moving both endpoints away from the surface

RFocus works best when either (or both) the transmitter and receiver is close to the surface. Performance degrades when both are moved away. Figure 14 shows that RFocus provides benefits up to a few meters from either transmitter or receiver. If ubiquitously deployed, such as on a wall, floor, or ceiling in every room, RFocus allows small devices to realize the benefits of a large antenna array. Alternately, surfaces could be placed close to APs or embedded in building fixtures.

7 Discussion

Efficiency In theory, an N-antenna phased array can achieve an $N\times$ gain in signal strength. RFocus achieves a median $10\times$

gain with 3200 antennas. RFocus suffers from two sources of loss in efficiency. First, only a fraction of the transmitted energy reaches the surface[4]. Second, RFocus can only direct a fraction of the energy that reaches it. This can be improved with better antenna design (§5.2) and metasurfaces [6].

RFocus's gain is qualitatively different from that of a 10 antenna phased array. RFocus precisely focuses a fraction of the transmitted energy, whereas the phased array coarsely directs all of the transmitted energy. Hence, near the receiver, RFocus's beam is more spatially concentrated while the phased array's coarse beam will interfere with nearby endpoints. Nevertheless, we believe these methods are complementary. Phased arrays can direct signals toward RFocus, which can then focus it precisely onto the receiver.

Wireless Controller In our prototype, elements are powered and controlled with wires. This isn't necessary. We envision a design where each element is powered and controlled wirelessly, like a passive RFID tag. The controller acts like a RFID controller that sets the state of each element. RFocus's fault tolerance allows deployment with failure-prone RFID tags (§6.4).

In such a setup, buildings could prefabricate their walls with RFocus elements. Carpets and wallpapers could be sold with RFocus elements already embedded in them. Users can separately buy a controller to control and obtain the benefits of the elements already present in the environment.

8 Conclusion

This paper presented RFocus, a system that moves beamforming functions from the radio transmitter to the environment. RFocus includes a two-dimensional surface with a rectangular array of simple elements, each of which contains an RF switch. Each element either lets the signal through or reflects it. The state of the elements is set by a software controller to maximize the signal strength at a receiver, using a majority-voting-based optimization algorithm. The RFocus surface can be manufactured as an inexpensive thin wallpaper, requiring no wiring. Our prototype implementation improves the median signal strength by $9.5\times$, and the median channel capacity by $2.0\times$.

Human Safety. Because RFocus doesn't emit any energy of its own, it does not increase the total radiation. It cannot focus energy to an area smaller than the size of a wavelength, since any device's ability to focus energy to an area smaller than a wavelength drops exponentially with distance from it. Hence RFocus is no riskier than being near the transmitter.

Ethics Statement. This work raises no ethical concerns.

---

[4]In most of our experiments, RFocus subtends $\approx 0.4\pi$ steradian on the receiver (equivalently, the transmitter). Hence it controls $\approx 10\%$ of the energy.

9 Acknowledgements

We thank Dinesh Bharadia, Peter Iannucci, Zach Kabelac, Dina Katabi, Colin Marcus, Vikram Nathan, Hariharan Rahul, Deepak Vasisht and Guo Zhang for useful discussion. We would also like to thank members of CSAIL's NetMIT group for letting us borrow their radio equipment.

References

[1] G. B. Airy. On the diffraction of an object-glass with circular aperture. *Trans. of the Cambridge Philosophical Society*, 5:283, 1835.

[2] I. Al-Naib and W. Withayachumnankul. Recent progress in terahertz metasurfaces. *Journal of Infrared, Millimeter, and Terahertz Waves*, 38(9):1067–1084, 2017.

[3] V. Arun and H. Balakrishnan. Rfocus: Practical beamforming for small devices. *arXiv preprint arXiv:1905.05130*, 2019.

[4] C. A. Balanis. Arrays: Linear, planar and circular. In *Antenna theory: analysis and design*, chapter 6. John wiley & sons, 2016.

[5] J. Chan, C. Zheng, and X. Zhou. WiPrint: 3D Printing Your Wireless Coverage. In *HotWireless*, 2015.

[6] H.-T. Chen, A. J. Taylor, and N. Yu. A review of metasurfaces: physics and applications. *Reports on progress in physics*, 79(7):076401, 2016.

[7] H.-T. Chen, A. J. Taylor, and N. Yu. A review of metasurfaces: physics and applications. *Reports on progress in physics*, 79(7):076401, 2016.

[8] M. Chen, M. Kim, A. M. Wong, and G. V. Eleftheriades. Huygens' metasurfaces from microwaves to optics: a review. *Nanophotonics*, 7(6):1207–1231, 2018.

[9] L. J. Chu. Physical limitations of omni-directional antennas. *Journal of applied physics*, 19(12):1163–1175, 1948.

[10] J. Costantine, Y. Tawk, S. E. Barbin, and C. G. Christodoulou. Reconfigurable antennas: Design and applications. *Proceedings of the IEEE*, 103(3):424–437, 2015.

[11] P. del Hougne, M. Fink, and G. Lerosey. Optimally diverse communication channels in disordered environments with tuned randomness. *Nature Electronics*, 2(1):36–41, 2019.

[12] M. Di Renzo and J. Song. Reflection probability in wireless networks with metasurface-coated environmental objects: An approach based on random spatial processes. *arXiv preprint arXiv:1901.01046*, 2019.

[13] P. Genevet and F. Capasso. Holographic optical metasurfaces: a review of current progress. *Reports on Progress in Physics*, 78(2):024401, 2015.

[14] J. Huang and J. A. Encinar. *Reflectarray antennas*, volume 30. John Wiley & Sons, 2007.

[15] N. Kaina, M. Dupré, G. Lerosey, and M. Fink. Shaping complex microwave fields in reverberating media with binary tunable metasurfaces. *Scientific reports*, 4(1):1–8, 2014.

[16] Z. Li, Y. Xie, L. Shangguan, R. I. Zelaya, J. Gummeson, W. Hu, and K. Jamieson. Towards programming the radio environment with large arrays of inexpensive antennas. In *NSDI*, 2019.

[17] C. Liaskos, S. Nie, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, and I. Akyildiz. A new wireless communication paradigm through software-controlled metasurfaces. *IEEE Communications Magazine*, 56(9):162–169, 2018.

[18] C. Liaskos, S. Nie, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, and I. Akyildiz. A novel communication paradigm for high capacity and security via programmable indoor wireless environments in next generation wireless systems. *Ad Hoc Networks*, 87:1–16, 2019.

[19] C. Liaskos, A. Tsioliaridou, S. Nie, A. Pitsillides, S. Ioannidis, and I. Akyildiz. Modeling, simulating and configuring programmable wireless environments for multi-user multi-objective networking. *arXiv preprint arXiv:1812.11429*, 2018.

[20] L. Lipson and H. Lipson. Optical physics. page 340, 1998.

[21] J. S. McLean. A re-examination of the fundamental limits on the radiation q of electrically small antennas. *IEEE Transactions on antennas and propagation*, 44(5):672, 1996.

[22] P. F. McManamon, P. J. Bos, M. J. Escuti, J. Heikenfeld, S. Serati, H. Xie, and E. A. Watson. A review of phased array steering for narrow-band electrooptical systems. *Proceedings of the IEEE*, 97(6):1078–1096, 2009.

[23] P. F. McManamon, P. J. Bos, M. J. Escuti, J. Heikenfeld, S. Serati, H. Xie, and E. A. Watson. A review of phased array steering for narrow-band electrooptical systems. *Proceedings of the IEEE*, 97(6):1078–1096, 2009.

[24] L. Novotny. *Lecture Notes on Electromagnetic Fields And Waves*. 2013.

[25] D. Parker and D. C. Zimmermann. Phased arrays–Part I: Theory and Architectures. *IEEE Trans. on Microwave Theory and Techniques*, 50(3):678–687, 2002.

[26] D. Parker and D. C. Zimmermann. Phased Arrays–Part II: Implementations, Applications, and Future Trends. *IEEE Trans. on Microwave Theory and Techniques*, 50(3):688–698, 2002.

[27] Y.-X. Ren, R.-D. Lu, and L. Gong. Tailoring light with a digital micromirror device. *Annalen der physik*, 527(7-8):447–470, 2015.

[28] L. Subrt and P. Pechac. Intelligent walls as autonomous parts of smart indoor environments. *IET Communications*, 6(8):1004–1010, 2012.

[29] X. Tan, Z. Sun, J. M. Jornet, and D. Pados. Increasing indoor spectrum sharing capacity using smart reflectarray. In *2016 IEEE International Conference on Communications (ICC)*, pages 1–6. IEEE, 2016.

[30] D. Tse and P. Viswanath. *Fundamentals of Wireless Communication*, chapter 2. Cambridge University Press, 2005.

[31] R. B. Waterhouse, S. Targonski, and D. Kokotoff. Design and performance of small printed antennas. *IEEE Transactions on Antennas and Propagation*, 46(11):1629–1633, 1998.

[32] A. Welkie, L. Shangguan, J. Gummeson, W. Hu, and K. Jamieson. Programmable radio environments for smart spaces. In *HotNets*, 2017.

[33] X. Xiong, J. Chan, E. Yu, N. Kumari, A. Sani, C. Zheng, and X. Zhou. Customizing indoor wireless coverage via 3d-fabricated reflectors. In *4th ACM International Conference on Systems for Energy-Efficient Built Environments*, 2017.

A Proof of Lemma 2

First, we prove another intermediate result.

Lemma 3. *Let $\mathbf{b}_\theta$ be the state such that $\mathbf{b}_\theta^{(i)} = 1$ iff $\Re(\mathbf{h}^{(i)} \cdot e^{-j\theta}) > 0$. If assumption 1 holds, then for any θ, the expected value of $\mathbf{h} \cdot \mathbf{b}_\theta$ is $\frac{e^{j\theta}}{\pi} \sum_i |\mathbf{h}^{(i)}|$.*

Proof.

$$E[\mathbf{h} \cdot \mathbf{b}_\theta] = \sum_i E[\mathbf{h}^{(i)} \mathbf{b}_\theta^{(i)}] = \sum_i E[e^{j\theta} \max \Re\{0, \mathbf{h}^{(i)} \cdot e^{-j\theta}\}]$$

$$= \frac{e^{j\theta}}{2} \sum_i E[\Re(\mathbf{h}^{(i)} \cdot e^{-j\theta})] = \frac{e^{j\theta}}{\pi} \sum_i |\mathbf{h}^{(i)}|$$

Here, $\max\Re$ compares the real component, and chooses $\mathbf{h}^{(i)} \cdot e^{-j\theta}$ if its real component is positive. The second step comes from our definition of $\mathbf{b}_\theta$. The third step notes that half of the values are expected to be non-zero and the sum of the remaining values is expected to have a phase of 0, and hence only the real part remains. For the final step, we compute the expected magnitude of the real part of a complex number with random phase in $(-\pi/2, \pi/2)$, given by the factor $\frac{1}{\pi} \int_{-\pi/2}^{\pi/2} \cos(\theta) d\theta = 2/\pi$. $\qquad\square$

A corollary of this is that, if $\phi = \text{Arg}(h_Z)$, $|E[h_Z + \mathbf{h} \cdot \mathbf{b}_\phi]| = |h_Z| + \frac{1}{\pi} \sum_i |\mathbf{h}^{(i)}|$. As discussed in §5.2, in an ideal system where we can choose any $\mathbf{b}^{(i)} \in \mathbb{C}, |\mathbf{b}^{(i)}| \leq 1$, the optimal assignment achieves a signal strength of $|h_Z| + \sum_i |\mathbf{h}^{(i)}|$. This gives us the same $1/\pi$ bound on what our system can achieve relative to the ideal when we account for h_Z and are willing to assume that phases are random (assumption 1).

Now, we restate and prove lemma 2 here.

Lemma 2. *Let $\mathbf{b}_{OPT}$ be the optimal assignment that maximizes $|h_Z + \mathbf{h} \cdot \mathbf{b}|$ and $\mathbf{b}_\perp$ be such that the ith component $\mathbf{b}_\perp^{(i)} = 1$ iff $\Re(\mathbf{h}^{(i)} \cdot h_Z^*) > 0$. If assumptions 1 and 3 hold, then as $N \to \infty$, $|H(\mathbf{b}_{OPT})|/|H(\mathbf{b}_\perp)| < 1 + \varepsilon \forall \varepsilon > 0$ with high probability.*

Proof. We will argue that $\mathbf{h} \cdot \mathbf{b}_\theta$ is roughly the same as the expected value (with high probability) for any θ. For any $\varepsilon > 0$ and $\theta, \phi \in (-\pi, \pi]$, let A_θ denote the event that $R_\theta = \frac{|\mathbf{h} \cdot \mathbf{b}_\theta|}{|E[\mathbf{h} \cdot \mathbf{b}_\phi]|} > 1 + \varepsilon$. We now show that for any $\varepsilon > 0$, the probability that $\frac{\text{Max}_{\theta \in (-\pi, \pi]} |\mathbf{h} \cdot \mathbf{b}_\theta|}{|E[\mathbf{h} \cdot \mathbf{b}_\phi]|} > 1 + \varepsilon$ goes to zero as $N \to 0$. Let p denote this

probability. Note, there are only N distinct values of $\mathbf{b}_\theta$ (say for $\theta_1, ..., \theta_N$), since it only changes when we include/exclude a new $\mathbf{h}^{(i)}$. Hence we need to take a maximum over only N values. Hence p is the probability that $\exists i$ such that that A_{θ_i} happens. That is, $p = P[\cup_{i=1}^N A_{\theta_i}] \leq \sum_{i=1}^N P[A_{\theta_i}] = NP[A_\theta]$, for any θ, since the expression is symmetric in θ (note: we need not assume A_{θ_i} are independent here). Let $\bar{h} = \sum_i |\mathbf{h}^{(i)}|/N$, be the mean of the magnitude individual components . Then $|E[\mathbf{h} \cdot \mathbf{b}_\theta]| = N\bar{h}/\pi$ and $R_\theta = \frac{\pi}{\bar{h}} |\sum_{i=1}^N \mathbf{h}^{(i)} \cdot \mathbf{b}_\theta^{(i)}|/N$. From the central limit theorem and lemma 3, we have

$$R_\theta = \frac{\pi}{\bar{h}} \left| \frac{\bar{h}}{\pi} + \mathcal{N}(0, \sigma^2/N) \right| = \left| 1 + \frac{\pi}{\bar{h}} \mathcal{N}(0, \sigma^2/N) \right|$$

where σ is the standard deviation in $\mathbf{b}_\theta^{(i)} \mathbf{h}^{(i)} \cos\theta$ for a random θ and $\mathcal{N}$ is a symmetric complex normal distribution with given mean and variance. Note, σ and $\bar{h}$ are bounded as $N \to \infty$, since assumption 3 tells us8 that $\mathbf{h}^{(i)}$ is bounded. Hence $P[A_\theta] = P\left[\mathcal{N}(0, \sigma^2/N) > \varepsilon \bar{h}/\pi\right]$ goes to zero as $N \to \infty$ for any $\varepsilon > 0$. $p = NP[A_\theta]$ also goes to zero as $N \to \infty$, since $P[A_\theta]$ decreases faster than $1/N$. We can follow a similar argument to show that for any $\varepsilon > 0$, the probability that $\frac{|\text{Min}_{\theta \in (-\pi, \pi]} \mathbf{h} \cdot \mathbf{b}_\theta|}{|E[\mathbf{h} \cdot \mathbf{b}_\theta]|} < 1 - \varepsilon$ goes to zero. And hence for any $\varepsilon' > 0$, the probability that $\frac{|E[\mathbf{h} \cdot \mathbf{b}_\theta]|}{|\text{Min}_{\theta \in (-\pi, \pi]} \mathbf{h} \cdot \mathbf{b}_\theta|} > 1 + \varepsilon'$ goes to zero.

Lemma 1 shows that $\exists \theta$ such that $\mathbf{b}_{OPT} = \mathbf{b}_\theta$. Hence

$$H(\mathbf{b}_{OPT}) = \text{Max}_\theta |h_Z + \mathbf{h} \cdot \mathbf{b}_\theta| \leq |h_Z| + |\text{Max}_\theta(\mathbf{h} \cdot \mathbf{b}_\theta)|$$

Also $\mathbf{b}_\perp = \mathbf{b}_\phi$ for $\phi = \text{Arg}(h_Z)$. Hence

$$H(\mathbf{b}_\perp) = |h_Z + \mathbf{h} \cdot \mathbf{b}_\phi| = |h_Z| + |\mathbf{h} \cdot \mathbf{b}_\phi| \geq |h_Z| + \text{Min}_\theta |\mathbf{h} \cdot \mathbf{b}_\theta|$$

The second step is true since the two terms have the same phase as $N \to \infty$. Thus,

$$\frac{H(\mathbf{b}_{OPT})}{H(\mathbf{b}_\perp)} \leq \frac{|h_Z| + \text{Max}_\theta |\mathbf{h} \cdot \mathbf{b}_\theta|}{|h_Z| + \text{Min}_\theta |\mathbf{h} \cdot \mathbf{b}_\theta|} \leq \frac{\text{Max}_\theta |\mathbf{h} \cdot \mathbf{b}_\theta|}{\text{Min}_\theta |\mathbf{h} \cdot \mathbf{b}_\theta|}$$

$$= \frac{\text{Max}_\theta |\mathbf{h} \cdot \mathbf{b}_\theta|}{E[|\mathbf{h} \cdot \mathbf{b}|]} \frac{E[|\mathbf{h} \cdot \mathbf{b}|]}{\text{Min}_\theta |\mathbf{h} \cdot \mathbf{b}_\theta|} < 1 + \varepsilon$$

with high probability for any $\varepsilon < 0$. This proves the theorem. $\qquad\square$