Randomized Roundings for a Mixed-Integer Elliptic Control System*

Martin T. Köhler[1], Lauri Kröger[1] and Christian Kirches[1]

Abstract— **We present randomized reconstruction approaches for optimal solutions to mixed-integer elliptic PDE control systems. Approximation properties and relations to sum-up rounding are derived using the cut norm. This enables us to dispose of space-filling curves required for sum-up rounding. Rates of almost sure convergence in the cut norm and the SUR norm in control space as well as almost sure H^1 convergence in state space are shown.**

I. INTRODUCTION

Elliptic control systems occur when analyzing stationary heat conduction, electrostatics, structural mechanics and many more. In some cases the control parameters are limited to a binary set. Such control problems exponentially increase in complexity as the discretization resolution increases. Solving a nonlinear integer program to optimality becomes practically infeasible.

Mixed-integer control of an elliptic system has been analyzed in depth by Manns with the use of sum-up rounding on a finite partition of the domain [1]. It has been shown that sum-up rounding has a weak-* convergence in the weak-* topology by strategically mapping the one-dimensional sum-up rounding algorithm to a two dimensional domain using space-filling curves.

We later show that in our context sum-up rounding generally does not converge. In turn we propose an algorithm which shows a stochastic convergence.

Contributions: We show an almost sure convergence of a Bernoulli rounding approach. Moreover, we demonstrate that sum-up rounding converges for randomized space-filling curves. Later on, we relate the convergence of the control variables, to the H^1 convergence of the state variables.

II. PRELIMINARIES

In this paper we consider the following mixed-integer control system on a bounded domain $\Omega := (0,1) \times (0,1)$:

$$\inf_{\bar{y}} \; J(\bar{y}) \text{ s.t. } L\bar{y} = \omega f \quad \text{in } \Omega, \tag{1a}$$

$$\bar{y} = g \quad \text{on } \partial\Omega, \tag{1b}$$

$$\omega \in \{0,1\} \quad \text{in } \Omega. \tag{1c}$$

Here, L denotes a linear second-order elliptic operator governing the state constraint (1a) and $\omega \in L^\infty(\Omega)$ represents a binary control variable for given $f \in L^\infty(\Omega)$ and boundary values g. We can further relax the binary constraint on ω by allowing it to take values in the interval $[0,1]$. This relaxation leads to the following formulation:

$$\min_{y} \; J(y) \text{ s.t. } Ly = \alpha f \quad \text{in } \Omega, \tag{2a}$$

$$y = g \quad \text{on } \partial\Omega, \tag{2b}$$

$$\alpha \in [0,1] \quad \text{in } \Omega. \tag{2c}$$

By regularly refining the cartesian grid of piecewise binary controls ω, obtained by strategically rounding the relaxed controls α, it has been demonstrated that the state variables under binary controls converge to those under the relaxed controls [1]. Finding a binary approximation to α is what we analyze in the context of the cut norm.

III. A BERNOULLI ROUNDING APPROACH

In this section, we introduce a new randomized approach to reconstruct a binary feasible control ω from a relaxed one denoted by α. We interpret the relaxed controls α as probabilities for the rounded controls ω to be 1, instead of 0, for almost all points $(x,y) \in \Omega$.

A. Bernoulli Rounding

Definition 1: Let $\left(\Omega_{(x,y)}, \mathcal{F}_{(x,y)}, \mathbb{P}_{(x,y)}\right)$ be a probability space, where the sample space is $\Omega_{(x,y)} := \{0,1\}$, the σ-algebra is $\mathcal{F}_{(x,y)} := \{\emptyset, \{0\}, \{1\}, \{0,1\}\}$, and the probability measure $\mathbb{P}_{(x,y)}$ is

$$\mathbb{P}_{(x,y)}(\omega(x,y) = k) := \begin{cases} 1 - \alpha(x,y) & \text{if } k = 0, \\ \alpha(x,y) & \text{if } k = 1. \end{cases} \tag{3}$$

The rounded control $\omega(x,y)$ follows a Bernoulli distribution with parameter $\alpha(x,y)$, denoted by $\omega(x,y) \sim \mathrm{Ber}(\alpha(x,y))$, for almost all $(x,y) \in \Omega$.

We define a sequence of partitions $\left\{P_i^{(n)}\right\}_{i=1}^n$ of $(0,1)$, where $P_i^{(n)} = \left(\frac{i-1}{n}, \frac{i}{n}\right)$ for $i = 1, \dots, n$. We choose ω piecewise constant on each set $P_i^{(n)} \times P_j^{(n)}$ such that

$$\omega_n(x,y) = \sum_{i=1}^n \sum_{j=1}^n \mathbb{1}_{P_i^{(n)}}(x) \mathbb{1}_{P_j^{(n)}}(y) \omega_{ij}, \tag{4}$$

where $\omega_{ij} \in \{0,1\}$ are the values that $\omega_n(x,y)$ take on the partition $P_i^{(n)} \times P_j^{(n)}$. Then ω_{ij} is a Bernoulli random variable with the parameter

$$\alpha_{ij} = \frac{1}{\left|P_i^{(n)}\right| \left|P_j^{(n)}\right|} \int_{P_i^{(n)} \times P_j^{(n)}} \alpha(x,y) \, \mathrm{d}x\mathrm{d}y. \tag{5}$$

The algorithm of Bernoulli Rounding (BR) summarizes as

$$\omega_{ij} \sim \mathrm{Ber}(\alpha_{ij}) \quad 1 \le i,j \le n. \tag{6}$$

Funded by the European Union. Views and opinions expressed are however those of the author(s) and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. This work is supported by ERC grant SCARCE, 101087662.

[1]Martin T. Köhler, Lauri Kröger, and Christian Kirches are with Institute for Mathematical Optimization, TU Braunschweig, Universitätsplatz 2, 38106 Braunschweig, GERMANY.
{martin.koehler1|lauri.kroeger1|c.kirches}@tu-bs.de

B. The Cut Norm

We now take interest in the difference $\|\omega - \alpha\|$ in a suitable norm. We make use of the *cut norm*, introduced in [2], which for an $M = (m_{ij}) \in \mathbb{R}^{n\times n}$ is given by

$$\|M\|_{\mathrm{C}} = \max_{S,T\subseteq[n]} \left|\sum\nolimits_{i\in S, j\in T} m_{ij}\right|. \tag{7}$$

As the node sets S and T are not required to be disjoint, the cut norm is an upper bound to the maximum cut capacity, such that the problem of computing $\|\cdot\|_{\mathrm{C}}$ is a relaxation of the max cut problem in matrix notation.

The cut norm proved useful in the context of graph limits, where sequences of finite graphs converge to a limit object called a graphon, with respect to the cut norm [3]. Graphs and graphons can be represented as kernel functions. The cut norm has therefore been generalized as follows:

Definition 2 (Cut Norm, cf. [3]): For a kernel $v \in L^1([0,1]\times[0,1])$, the cut norm is defined by

$$\|v\|_{\square} = \sup_{S,T\subseteq[0,1]} \left|\int_{S\times T} v(x,y)\,\mathrm{d}x\mathrm{d}y\right| \tag{8}$$

for measurable subsets S and T.

After introducing piecewise constant elements for ω and α, the (normalized) cut norm becomes

$$\|\omega-\alpha\|_{\square} = \frac{1}{n^2}\|\omega-\alpha\|_{\mathrm{C}} = \max_{S,T\subseteq[n]} \frac{1}{n^2}\Big|\sum_{i\in S,j\in T}\omega_{ij}-\alpha_{ij}\Big|, \tag{9}$$

and will be denoted by $\|\cdot\|_{\square}$ to distinguish it from its non-normalized counterpart $\|\cdot\|_{\mathrm{C}}$. The normalization factor $1/n^2$ represents the volume of a discretization element. The following lower bound [4] will be required later:

$$\left|\int_{[0,1]\times[0,1]} \omega(x,y)-\alpha(x,y)\,\mathrm{d}x\mathrm{d}y\right| \le \|\omega-\alpha\|_{\square}. \tag{10}$$

C. Almost Sure Convergence

We are now prepared to state the first central theorem of this contribution.

Theorem 1: Let $\alpha : (0,1)\times(0,1) \to [0,1]$ be a measurable function. Consider $\omega_n : (0,1)\times(0,1) \to \{0,1\}$ to be a piecewise constant function on the product sets from the sequence of partitions $\left\{P_i^{(n)}\right\}_{i=1}^n$. If $\omega_n(x,y) = \omega_{ij}$ is a Bernoulli random variable for $(x,y) \in P_i^{(n)} \times P_j^{(n)}$ with parameter α_{ij} as in (5), then it follows that

$$\lim_{n\to\infty} \|\omega_n - \alpha\|_{\square} = 0 \quad \text{almost surely.} \tag{11}$$

Note that this statement is not immediate from the cut norm definition, as the maximizing choice of sets S, T may always select on the order of n^2 elements from the matrix, thus fully compensating the weight $1/n^2$.

The relation of the cut norm to the operator norm and a result from random matrix theory will be of use in this proof.

Proof: Define $\nu := \omega_n - \alpha$ with $\omega_n, \alpha \in \mathbb{R}^{n\times n}$ and note that the elements are in $[-1,1]$ and have mean zero. Using Lemma 2 (a random matrix inequality) from the appendix, we have

$$\mathbb{P}\left(\sqrt{n}\,\|\nu x\|_2 \ge An\right) \le C\exp(-cA^2 n), \tag{12}$$

with the restriction $\|x\|_\infty = 1$ and constants $A, C, c > 0$. We now make use of the cut norm bound [2]

$$\|\nu\|_C \le \sup_{x\in\mathbb{R}^n\setminus\{0\}} \frac{\|\nu x\|_1}{\|x\|_\infty} \le \sup_{x\in\mathbb{R}^n\setminus\{0\}} \frac{\sqrt{n}\,\|\nu x\|_2}{\|x\|_\infty}, \tag{13}$$

and obtain

$$\mathbb{P}\left(\|\nu\|_{\mathrm{C}} \ge An\right) \le C\exp(-cA^2 n). \tag{14}$$

Let $A = sn$ represent the cut norm being greater than some $s > 0$, i.e. $1/n^2\,\|\cdot\|_{\mathrm{C}} \ge s$. The constant A is certain to become sufficiently large eventually, and we see that

$$\mathbb{P}(\|\nu\|_{\mathrm{C}} \ge sn^2) \le C\exp(-cs^2n^3). \tag{15}$$

Defining events $E_n := \{\|\nu\|_{\mathrm{C}} \ge sn^2\}$, we consider the sums

$$\sum_{n=1}^{\infty} \mathbb{P}(E_n) \le \sum_{n=1}^{\infty} C\exp(-cs^2n^3). \tag{16}$$

This series converges as $n \to \infty$, therefore is finite. Following Borel-Cantelli [5]

$$\mathbb{P}\left(\limsup_{n\to\infty} E_n\right) = 0. \tag{17}$$

That is, the event that the cut norm becomes greater than some $s > 0$ occurs only finitely many times as $n \to \infty$. Therefore, the probability that the Bernoulli approach yields a cut norm greater than some $s > 0$ goes to zero almost surely as we refine the discretization grid. ■

IV. SUM-UP ROUNDING

In this section, we relate the Bernoulli rounding algorithm (BR) to the established sum-up rounding algorithm. Sum-up rounding (SUR) was introduced by [6], [7] for mixed-integer ODE constrained optimal control problems (MIOCPs). Its purpose is to construct binary feasible controls that are provably close to relaxed optimal ones. Of relevance to our setting is the problem

$$\inf_{\omega}\ J(\omega) \text{ s.t. } \dot{x}(t) = A(x(t))\omega(t) \quad \text{a.e. in } (0,t_f], \tag{18a}$$
$$x(0) = x_0, \tag{18b}$$
$$\omega \in \{0,1\} \quad \text{a.e. in } [0,t_f]. \tag{18c}$$

The approximation approach is to introduce relaxed control variables $\alpha(t) \in [0,1]$ and to solve the problem using the differential equation constraint $\dot{x}(t) = A(x(t))\alpha(t)$. Afterwards, a relaxed optimal control α is rounded on a given time grid $0 = t_0 < t_1 < \cdots < t_m = t_f$ such that the distance measure $d_{\mathrm{SUR}}(\alpha,\omega) = \|\alpha-\omega\|_{\mathrm{SUR}}$ implied by the *SUR norm*

$$\|\nu\|_{\mathrm{SUR}} := \sup_{t\in[0,t_f]} \left|\int_0^t \nu(\tau)\,\mathrm{d}\tau\right| \tag{19}$$

becomes small.

To this end, assume ω and α to be piecewise constant on $[0,t_f]$, that is, $\omega_i := \omega(t)$ and $\alpha_i := \alpha(t)$ for all $t \in [t_i, t_{i+1}]$.

The *sum-up rounding algorithm* reconstructs a binary control ω from a relaxed one α by computing ($\Delta t_i = t_{i+1} - t_i$)

$$\omega_i = \begin{cases} 1 & \text{if } \sum_{k=0}^{i-1}(\alpha_k - \omega_k)\Delta t_k + \alpha_i \Delta t_i \geq \frac{1}{2}\Delta t_i, \\ 0 & \text{otherwise.} \end{cases} \tag{20}$$

Fundamentally, SUR is a one-dimensional algorithm, and its convergence in the control and state space are well understood on interval domains, cf. [7], [8]. Manns et al. [1] proposed mapping two-dimensional domains using space filling curves $c : [0, 1] \to [0, 1] \times [0, 1]$. The approach is to dissect the closure of the domain Ω into a sequence of cells $\left(\{Q_1^{(n)}, \ldots, Q_{n^2}^{(n)}\}\right)_n$ such that $\cup_{i=1}^{n^2} Q_i^{(n)} = \bar{\Omega}$ for all $n \in \mathbb{N}$. On each partition, define $\omega_i = \frac{1}{n^2}\int_{Q_i^{(n)}} \omega \,\mathrm{d}x\mathrm{d}y$. A binary control ω is then computed from α by carrying out SUR along the cells on the curve, and the curve-dependent SUR norm is

$$\|\nu\|_{\text{SUR}}^c = \sup_{t\in[0,1]} \left| \int_0^t \nu(c(\tau))\,\mathrm{d}\tau \right|. \tag{21}$$

Weak-* convergence $\sum_{i=1}^{n^2} \omega \mathbb{1}_{Q_i^{(n)}} \overset{*}{\rightharpoonup} \alpha$ as $n \to \infty$ of SUR along space-filling curves and a rate of convergence in $\|\cdot\|_{H^{-1}(\Omega)}$ have been shown in [1].

A. SUR Converges a.s. in the Cut Norm

Like BR, SUR also converges almost surely in the cut norm. This result however requires a slightly different argument. We first consider a worst-case example given by the relaxed controls $\alpha \equiv 1/2$ on a discretized domain represented by the following $n \times n$ matrix

$$\alpha = \begin{pmatrix} 1/2 & \cdots & 1/2 \\ \vdots & \ddots & \vdots \\ 1/2 & \cdots & 1/2 \end{pmatrix} \tag{22}$$

with n even. We traverse the domain following any curve obeying the von Neumann neighborhood, and use SUR to generate a binary feasible control, which will always yield

$$\omega = \begin{pmatrix} 1 & 0 & 1 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 1 \\ 1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & 0 & \cdots & 1 \end{pmatrix}, \ \nu = \begin{pmatrix} +1/2 & -1/2 & +1/2 & \cdots & -1/2 \\ -1/2 & +1/2 & -1/2 & \cdots & +1/2 \\ +1/2 & -1/2 & +1/2 & \cdots & -1/2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1/2 & +1/2 & -1/2 & \ldots & +1/2 \end{pmatrix}. \tag{23}$$

We notice that throughout the domain we obtain a checkerboard pattern of zeros and ones. Considering the difference $\nu = \omega - \alpha$ in the normalized cut norm (9), this always leads to either $S = T = \{1, 3, 5, \ldots, n-1\}$ or $S = T = \{2, 4, 6, \ldots, n\}$. Thus, the value of the cut norm is, independent of n,

$$\|\nu\|_\square = \frac{1}{n^2}|S||T|\frac{1}{2} = \frac{1}{n^2} \cdot \frac{n}{2} \cdot \frac{n}{2} \cdot \frac{1}{2} = \frac{1}{8}. \tag{24}$$

One has the bound [2]

$$\sup_{x\in\mathbb{R}^n\setminus\{0\}} \frac{\|\nu x\|_1}{\|x\|_\infty} \leq 4\|\nu\|_{\text{C}} \tag{25}$$

and can see the operator norm to equal $n^2/2$ (one lets $x = (1, -1, 1, -1, \ldots)^\top$, which is extremal since the objective is convex and the feasible set a polytope). Hence this example constitutes the worst case.

On the other hand, the situation turns out to be an exception triggered by the curve we followed when computing the SUR norm. For random curves, the situation is almost surely better. We keep in mind the bound (13), which relates the cut norm to the operator norm [2].

Now consider matrix ν to be a random matrix with i.i.d. entries $\nu_{ij} \in \{-1/2, +1/2\}$, both with equal probability $1/2$. Then, elementary random matrix theory knows the following bound (Lemma 2 in the appendix):

$$\begin{aligned} &\mathbb{P}(2\|\nu x\|_2 \geq A\sqrt{n}) \leq C\exp(-cA^2 n) \\ \iff &\mathbb{P}(\sqrt{n}\|\nu x\|_2 \geq An/2) \leq C\exp(-cA^2 n). \end{aligned} \tag{26}$$

The argumentation is analogous to that of the proof of Theorem 1 and from that it follows

$$\mathbb{P}(\|\nu\|_{\text{C}} \geq sn^2) \leq C\exp(-4cs^2 n^3) \tag{27}$$

and an almost sure convergence as $n \to \infty$.

Now, the choice of a random matrix ν can be interpreted as sampling with replacement from the $\omega - \alpha$ instance considered above. Carrying out SUR along a random curve however amounts to sampling *without* replacement, that is, random shuffles of the elements of matrix ν. The cut norm is maximized if S and T can be chosen such that the subset of elements selected contains as many $+1/2$ elements (or as many $-1/2$ elements) as possible. Under sampling without replacement, the budget of such elements is restricted to exactly $n/2$ many, such that attaining higher cut norms becomes less likely. Hence, the bound just proved also holds for sampling without replacement, i.e., for random curves.

B. BR Converges a.s. in the SUR Norm Along Any Curve

We can show that our Bernoulli rounding approach does converge almost surely in the norm $\|\cdot\|_{\text{SUR}}^\pi$. On a discretized domain, we can represent the norm as

$$\|\omega - \alpha\|_{\text{SUR}}^\pi = \max_{k\in[n^2]} \frac{1}{n^2} \left| \sum_{i=1}^k \omega_{\pi(i)} - \alpha_{\pi(i)} \right|, \tag{28}$$

where π encodes the curve as a permutation of cells $Q_1^{(n)}, \ldots, Q_{n^2}^{(n)}$.

Theorem 2: Let the assumptions of Theorem 1 for α and the piecewise constant function ω_n hold, then

$$\lim_{n\to\infty} \|\omega_n - \alpha\|_{\text{SUR}}^\pi = 0 \quad \text{almost surely.} \tag{29}$$

Proof: We insert the sum of the SUR norm into the inequality of Lemma 1 (a Hoeffding inequality),

$$\mathbb{P}\left(\left|\sum_{i=1}^k \omega_{\pi(i)} - \alpha_{\pi(i)}\right| \geq t\right) \leq 2\exp\left(-\frac{2t^2}{k}\right) \tag{30}$$

where index $1 \leq k \leq n^2$ indicates the cell $Q_{\pi(k)}^{(n)}$ that gave rise to the supremum in the SUR norm. Now, we relate t

to the SUR norm by defining $t := sn^2$, representing the difference in the norm being greater than some $s > 0$, i.e.

$$\mathbb{P}\left(\left|\sum_{i=1}^{k}\omega_{\pi(i)} - \alpha_{\pi(i)}\right| \geq sn^2\right) \leq 2\exp\left(-\frac{2(sn^2)^2}{k}\right)$$
$$\leq 2\exp\left(-\frac{2s^2n^4}{n^2}\right) = 2\exp\left(-2s^2n^2\right). \quad (31)$$

To continue, we refer to the proof of Theorem 1 as the rest is analogous. An almost sure convergence as $n \to \infty$ follows from the argument of Borel-Cantelli [5]. ∎

The almost sure convergence of the Bernoulli Rounding approach in the SUR norm is independent of the permutation π, i.e., of the curve used to compute the SUR norm.

C. Convergence Rates of BR in Both Norms

In the probability bound for BR in the cut norm (15), we can choose $s^2 \in o(n^3)$ to ensure almost sure convergence and similarly, for the SUR norm in (31), we will select $s^2 \in o(n^2)$ to guarantee convergence. Specifically, this implies that $s \in o(n^{1.5})$ for the cut norm and $s \in o(n^1)$ for the SUR norm.

Both norms intrinsically decrease at the rate of $\mathcal{O}(n^2)$, and the convergence rates are determined by the products of this rate with the respective rates for s. For the cut norm, the product $\mathcal{O}(n^2) \cdot \mathcal{O}(n^{-1.5}) = \mathcal{O}(n^{0.5})$ dictates the convergence rate for BR in the cut norm and the product $\mathcal{O}(n^2) \cdot \mathcal{O}(n^{-1}) = \mathcal{O}(n^1)$ determines the convergence rate in the SUR norm respectively.

V. STATE CONVERGENCE

Of the cut norm and the SUR norm, neither one can be shown to dominate the other. Hence, in this section, we show that cut norm convergence in the space of reconstructed binary controls induces norm converges in the space of differential states.

We assume $\alpha f \in L^\infty(\Omega)$ and appropriate boundary conditions g, such that there exist unique weak solutions $y \in H^1(\Omega)$ [9].

The linear operator L has the general divergence form

$$Ly = -\sum_{i,j=1}^{n} a^{ij}(x) y_{x_i x_j} + \sum_{i=1}^{n} b^i(x) y_{x_i} + c(x)y \quad (32)$$

where in our scenario $n = 2$ because $\Omega \subset \mathbb{R}^2$. Furthermore, assume $a^{ij}, b^i, c \in L^\infty(\Omega)$ and the symmetry $a^{ij} = a^{ji}$.

We continue with the weak formulation of (2a) and define the bilinear form $a(\cdot,\cdot) : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ to be

$$a(y, v) := -\int_\Omega \sum_{i,j=1}^{n} a^{ij} y_{x_i} v_{x_j} + \sum_{i=1}^{n} b^i y_{x_i} v + cyv \,\mathrm{d}x \quad (33)$$

and

$$\langle \alpha f, v\rangle := \int_\Omega \alpha f v \,\mathrm{d}x, \quad (34)$$

where $\langle\cdot,\cdot\rangle$ denotes the inner product in $L^2(\Omega)$. We are now concerned with

$$a(y, v) = \langle \alpha f, v\rangle. \quad (35)$$

Let y be the solution under relaxed controls and $\bar{y}$ the solution of the binary controls. Consider the difference $w := \bar{y} - y$ such that

$$Lw = (\omega - \alpha) f \quad \text{in } \Omega, \quad (36a)$$
$$w = 0 \quad \text{on } \partial\Omega. \quad (36b)$$

Because of the homogeneous boundary condition, weak solutions w are in the space $H_0^1(\Omega)$.

We know that w is bounded in $L^\infty(\Omega)$ for a given $(\omega - \alpha) f$ [9] and there exists a constant $\tilde{C} > 0$ such that

$$\|w\|_{L^\infty(\Omega)} \leq \tilde{C}\, \|(\omega - \alpha) f\|_{L^\infty(\Omega)}. \quad (37)$$

The constant $\tilde{C}$ depends only on the diameter of Ω. From this relation, however, we cannot deduce, that if $\omega \to \alpha$ with respect to the cut norm, $w \to 0$, but the bound will be useful in the following proof.

Theorem 3: Assume $f \geq 0$. If $\|\omega - \alpha\|_\square \to 0$, then $\|w\|_{H_0^1(\Omega)} \to 0$.

Proof: Since L is a second-order elliptic operator, there exists a constant $C > 0$ such that

$$C\,\|w\|_{H_0^1(\Omega)} \leq a(w, w) \quad (38)$$

for any weak solution $w \in H_0^1(\Omega)$. Additionally using the Hölder inequality and the lower bound to the cut norm (10), we can estimate

$$\begin{aligned}\langle(\omega - \alpha) f, v\rangle &\leq \operatorname*{ess\,sup}_\Omega (fv) \int_\Omega \omega - \alpha \,\mathrm{d}x \\ &\leq \operatorname*{ess\,sup}_\Omega (fv)\, \|\omega - \alpha\|_\square\end{aligned} \quad (39)$$

for all $v \in H_0^1(\Omega)$. Substituting $v = w$, and noting $\operatorname{ess\,sup}_\Omega(fw)$ is bounded because of (37), we obtain

$$\|w\|_{H_0^1(\Omega)} \leq \frac{1}{C} \operatorname*{ess\,sup}_\Omega (fw)\, \|\omega - \alpha\|_\square. \quad (40)$$

Therefore, as $\|\omega - \alpha\|_\square \to 0$, it follows that $\|w\|_{H_0^1(\Omega)} \to 0$. ∎

As the BR and SUR algorithms converge almost surely in the control space with respect to the cut norm, we know that an H^1 convergence in the state space follows.

VI. NUMERICAL RESULTS

In this section, we present results of numerical experiments. First, we consider the Poisson equation with a control variable on the right-hand side

$$-\Delta y = u \quad \text{in } \Omega := (0, 1) \times (0, 1), \quad (41a)$$
$$y = 0 \quad \text{on } \partial\Omega. \quad (41b)$$

The problem is discretized using a finite element scheme in FEniCS [10] with a Friedrich-Keller triangulation for the state and piecewise constant functions on the respective quadrilateral grid for the control variables [11]. The goal is to investigate the changes in the state variables, when we reconstruct a binary control that approximates u in the cut norm. As a benchmark, we set $u \equiv 0.5$ in Ω and compare the resulting states in the L^2 and H^1 norm. Since the controls are piecewise constant, we can easily map them to a matrix

$U, \bar{U} \in \mathbb{R}^{n \times n}$, where $\bar{U} \in \{0,1\}^{n \times n}$ represents the binary controls.

A. Computing the Cut Norm

To compute S, T in the cut norm, we need to solve the following bipartite binary quadratic optimization problem:

$$\max_{s,t} \left| t^\top (\bar{U} - U) s \right| \text{ s.t. } s, t \in \{0,1\}^n. \tag{42}$$

The quadratic form is indefinite and the problem of computing the global maximizer that gives rise to the cut norm objective is hard. In [12], MaxSNP-hardness is shown and a PTIME $\varepsilon-$approximation algorithm with error bound εN^2 is offered, but its implementation is not without effort.

If instead one solves the relaxed quadratic optimization problem twice in standard form with $\pm Q := \pm(\bar{U} - U)$,

$$\max_{s,t} \begin{pmatrix} s \\ t \end{pmatrix}^\top \begin{pmatrix} 0 & Q^\top \\ Q & 0 \end{pmatrix} \begin{pmatrix} s \\ t \end{pmatrix} \text{ s.t. } 0 \le s, t \le 1. \tag{43}$$

one observes that all stationary points of the relaxation are immediately binary feasible. With duals $\lambda_s, \lambda_t \in \mathbb{R}^n$, we find the stationarity conditions

$$Q^\top t = [\lambda_s]^- + [\lambda_s]^+(1 - s), \; Qs = [\lambda_t]^- + [\lambda_t]^+(1 - t).$$

A fractional solution component $s_i \in (0,1)$ would imply $(Q^\top t)_i = (\lambda_s)_i = 0$ due to complementary slackness, hence the objective contribution $s_i(Q^\top t)_i$ is zero. Either reducing or increasing the fractional s_i will generate a positive objective contribution for the same vector t. We hence rely on solving the non-convex relaxation using the global optimizer EAGO.jl [13]. It quickly finds very good upper bounds, but typically takes very long to confirm optimality. We therefore terminate early after 100 iterations, and fractional results are rounded. While this comes at the peril of underestimating the true cut norm, we have not observed adverse effects.

B. Results

In Figure 1, $\|\bar{y} - y\|$ is shown for the L^2 and H^1 norm of the problem (41). For each resolution we have applied BR and SUR reconstructions $\bar{u}$ of relaxed controls $u \equiv 0.5$ on an $n \times n$ grid and we display the mean and standard deviation over 100 samples. We compare the convergence to SUR along a Hilbert space-filling curve of which the convergence rate is known [1], [14]. We see that SUR along a randomized space-filling curve converges in both the L^2 and H^1 norm, reflecting the results in [1] for weaker assumptions on the space-filling curve. Our Bernoulli rounding approach converges with the same rate as SUR along a random curve. In Figure 3, we visualize the control and state variables and notice, that for higher resolutions, the state is visually indistinguishable from the one with relaxed controls.
Figure 2 shows the mean distance $\|\bar{u} - u\|$ in the cut and SUR norm over more than 100 samples. The convergence rates of BR and SUR in the cut and SUR norm correspond to the rates shown theoretically. Even though BR and SUR converge at the same rate in the cut norm, SUR is slightly better on average.

Furthermore, in Figure 4, we applied the rounding approaches to an optimal control problem. There, the cost function states $J(y) = \int_\Omega (y - y_d)^2 \mathrm{d}x + \int_\omega u \, \mathrm{d}x$ for given $y_d = 1/(2\pi^2) \sin(\pi x_1) \sin(\pi x_2)$, with zero Dirichlet boundaries only on the bottom and right edge of the domain and zero Neumann boundaries on the remaining edges. The optimal solution is computed using dolfin-adjoint [15].

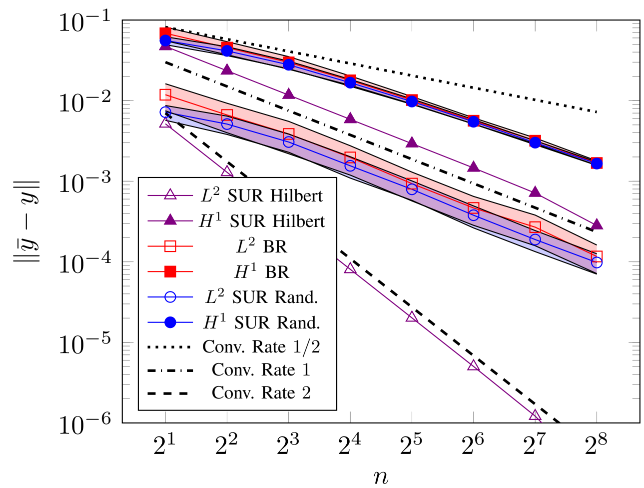

Fig. 1. Mean and one standard deviation of $\|\bar{y} - y\|$ in the L^2 and H^1 norm, with $u \equiv 0.5$ and reconstructed binary controls using BR and using SUR along a Hilbert curve and random curves

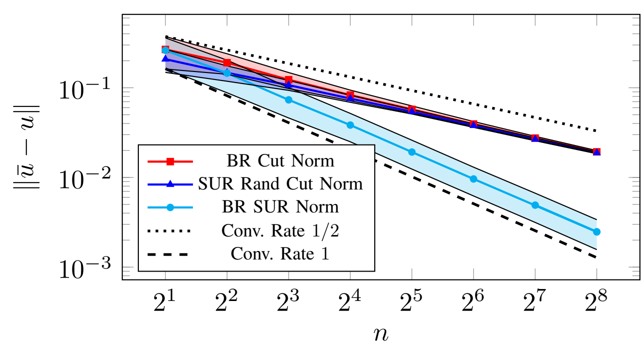

Fig. 2. Mean and one standard deviation of $\|\bar{u} - u\|$ in the cut and SUR norm, with $u \equiv 0.5$ and binary reconstructions $\bar{u}$ using BR and SUR along random curves

VII. Conclusion

We have shown that a simple Bernoulli rounding approach converges almost surely not only in the cut norm, but also in the SUR norm, related to the established SUR algorithm. For the SUR algorithm, we demonstrated that along a randomized space-filling curve, it converges almost surely in the cut norm. We then related the convergence of the control variables in the cut norm to the H^1 convergence in the state space.

Appendix

Lemma 1 (Hoeffding Inequality, cf. [16]): Let $X_1, X_2, \dots, X_r$ be independent random variables, each

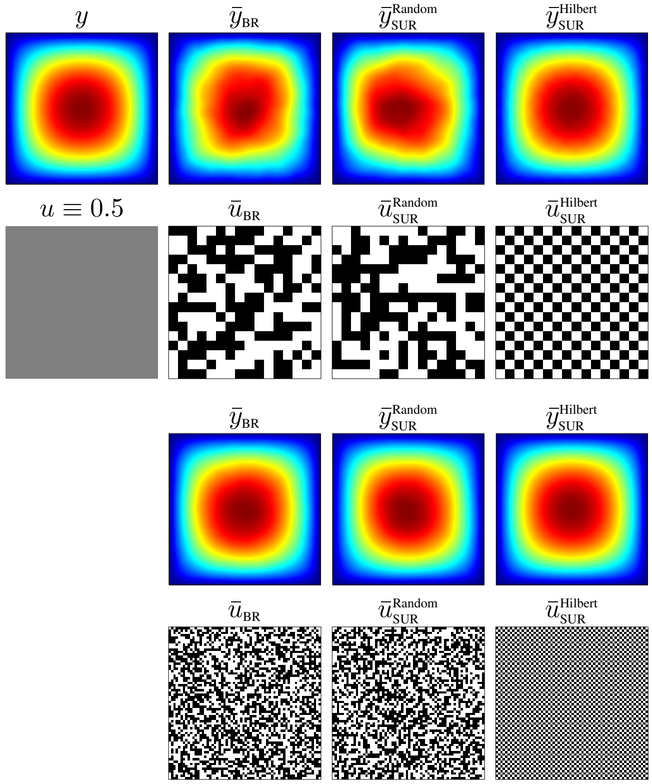

Fig. 3. State (top) and control variables (bottom), first control variable is 0.5, the rest show binary reconstructions on a 16×16 and 64×64 grid with BR and SUR along a random curve and along the Hilbert curve

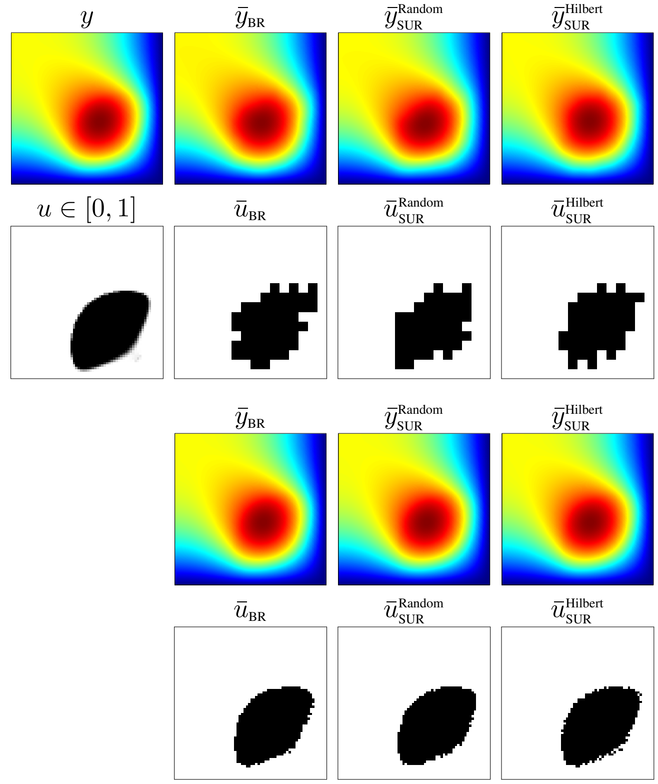

Fig. 4. State (top) and control variables (bottom), first control variable is relaxed, the rest show binary reconstructions on a 16×16 and 64×64 grid with BR and SUR along a random curve and along the Hilbert curve

$0 \le X_i \le 1$ for $i = 1, \dots, r$. Let $S = X_1 + \cdots + X_r$, then for all $t > 0$

$$\mathbb{P}\left(|S - \mathbb{E}(S)| \ge t\right) \le 2 \exp\left(-\frac{2t^2}{r}\right) \tag{44}$$

with $\mathbb{E}(S)$ the expected value.
A proof can be found in, e.g., [16].

Lemma 2 (Operator norm of a random matrix, cf. [17]): Let the coefficients of matrix $M = (m_{ij}) \in \mathbb{R}^{n \times n}$ be independent random variables with zero mean and uniformly bounded in magnitude by 1. Let x be a unit vector in $\mathbb{R}^n$. Then for A larger than some absolute constant, one has

$$\mathbb{P}(\|Mx\|_2 \ge A\sqrt{n}) \le C \exp(-cA^2 n) \tag{45}$$

for some absolute constants C, $c > 0$.
A proof can be found in, e.g. [17] with corrections available online[1].

References

[1] P. Manns and C. Kirches, "Multidimensional sum-up rounding for elliptic control systems," *SIAM Journal on Numerical Analysis*, vol. 58, no. 6, pp. 3427–3447, 2020.

[2] A. Frieze and R. Kannan, "Quick approximation to matrices and applications," *Combinatorica*, vol. 19, no. 2, pp. 175–220, 1999.

[3] L. Lovász, *Large Networks and Graph Limits*, ser. American Mathematical Society colloquium publications. American Mathematical Society, 2012.

[1]URL: https://terrytao.wordpress.com/2010/01/09/254a-notes-3-the-operator-norm-of-a-random-matrix/

[4] S. Janson, "Graphons, cut norm and distance, couplings and rearrangements," *NYJM Monographs*, vol. 4, 2013.

[5] T. Chandra, *The Borel-Cantelli Lemma*, ser. SpringerBriefs in Statistics. Springer India, 2012.

[6] S. Sager, *Numerical methods for mixed-integer optimal control problems*. Tönning: Der Andere Verlag, 2005.

[7] S. Sager, H. Bock, and M. Diehl, "The integer approximation error in mixed-integer optimal control," *Mathematical Programming*, vol. 133, no. 1, pp. 1–23, 2012.

[8] C. Kirches, F. Lenders, and P. Manns, "Approximation properties and tight bounds for constrained mixed-integer optimal control," *SIAM J. Control and Optimization*, vol. 58, no. 3, pp. 1371–1402, 2020.

[9] L. Evans, *Partial Differential Equations*, ser. Graduate studies in mathematics. American Mathematical Society, 1998.

[10] M. S. Alnaes, J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. Logg, C. N. Richardson, J. Ring, M. E. Rognes, and G. N. Wells, "The FEniCS project version 1.5," *Arch. Numerical Software*, vol. 3, 2015.

[11] F. Tröltzsch, *Optimal Control of Partial Differential Equations: Theory, Methods, and Applications*, ser. Graduate studies in mathematics. American Mathematical Society, 2010.

[12] N. Alon and A. Naor, "Approximating the Cut-Norm via Grothendieck's inequality," *STOC '04: Proceedings of the thirty-sixth annual ACM symposium on Theory of computing*, pp. 72–80, 2004.

[13] M. E. Wilhelm and M. D. Stuber, "Eago.jl: easy advanced global optimization in julia," *Optimization Methods and Software*, vol. 37, no. 2, pp. 425–450, 2022.

[14] C. Kirches, P. Manns, and S. Ulbrich, "Compactness and convergence rates in the combinatorial integral approximation decomposition," *Mathematical Programming*, vol. 188, no. 2, pp. 569–598, 2021.

[15] S. K. Mitusch, S. W. Funke, and J. S. Dokken, "dolfin-adjoint 2018.1," *Journal of Open Source Software*, vol. 38, no. 4, 2018.

[16] W. Hoeffding, "Probability inequalities for sums of bounded random variables," *J. Amer. Stat. Assoc.*, vol. 58, no. 301, pp. 13–30, 1963.

[17] T. Tao, *Topics in Random Matrix Theory*, ser. Graduate studies in mathematics. American Mathematical Society, 2012.