Finite-Sample Error Bound

I. Finite-sample error bound

The following result gives high-probability finite-sample error bounds for the average-value-at-risk constraint given a solution of the sample-based approximation $\textbf{SOCP}_M(\bar{\omega})$ of **OCP**. For conciseness, we use notations from the main paper that we do not reintroduce.

Lemma 1 (Error bound for solutions of $\textbf{SOCP}_M(\bar{\omega})$): Let $M \in \mathbb{N}$ be a number of samples, $\{\omega^i\}_{i=1}^M$ be independent and identically distributed (iid) samples of ω, and formulate the sample average approximation $\textbf{SOCP}_M(\bar{\omega})$ of **OCP**. Denote by $u_M(\bar{\omega})$ the solution of $\textbf{SOCP}_M(\bar{\omega})$. Then, under assumptions **(A1-A4)** (see the main paper), for some finite constants $(\tilde{C}, \bar{h})$ and a compact set $\mathcal{T} \subset \mathbb{R}$ large-enough, for (ϵ_M, δ_M) defined in (5), with probability at least $1 - \delta_M$ over the M iid samples ω^i,

$$\text{AV@R}_\alpha\left(\sup_{s\in[0,T]} G(x_{u(\bar{\omega})}(s), \xi)\right) \leq \epsilon_M. \tag{1}$$

Lemma 1 gives a finite-sample high-probability error bound for the satisfaction of the risk constraint in **OCP**. Specifically, replacing (16b) in $\textbf{SOCP}_M(\bar{\omega})$ (see the main paper) with

$$t + \frac{1}{\alpha M}\sum_{i=1}^M \max\left(\sup_{s\in[0,T]} G(x_u^i(s), \xi^i) - t, 0\right) \leq -\epsilon_M$$

suffices to guarantee the satisfaction of the AV@R_α constraint in **OCP**. Note that $\epsilon_M \to 0$ and $\delta_M \to 0$ as $M \to \infty$, i.e., the error can be made arbitrarily small (with increasingly high probability) by increasing the number of samples M. It should not be surprising to the reader that no additional assumptions are required to derive the finite-sample error bound in Lemma 1, as it relies on concentration inequalities used to prove Theorem 1 in the main paper. The remainder of this document describes the proof of Lemma 1.

II. Expectations concentration inequalities

The proof of Lemma 1 relies on a concentration inequality from [1]. Let $(\Omega, \mathcal{G}, \mathbb{P})$ be a probability space, $d \in \mathbb{N}$, $\mathcal{U} \subset \mathbb{R}^d$ be a compact set, $h : \mathcal{U} \times \Omega \to \mathbb{R}$ be a Carathéodory function (i.e., $h(u, \cdot)$ is $\mathcal{G}$-measurable for all $u \in \mathcal{U}$ and $h(\cdot, \omega)$ is continuous $\mathbb{P}$-almost-surely). We introduce the following two assumptions on h.

(B1) $\mathbb{P}$-almost-surely, the map $u \mapsto h(u, \omega)$ is α-Hölder continuous for some exponent $\alpha \in (0, 1]$ and Hölder constant $M(\omega)$ satisfying $\mathbb{E}[M(\cdot)^2] < \infty$, such that

$$|h(u_1, \omega) - h(u_2, \omega)| \leq M(\omega)\|u_1 - u_2\|_2^\alpha \tag{2}$$

for all $u_1, u_2 \in \mathcal{U}$.

(B2) For some constant $0 \leq \bar{h} < \infty$, $\mathbb{P}$-almost-surely,

$$\sup_{u\in\mathcal{U}} |h(u, \omega)| \leq \bar{h}. \tag{3}$$

Let $\{\omega^i\}_{i=1}^M$ be $M \in \mathbb{N}$ independent and identically distributed (iid) samples of ω. We have the following result.

Lemma 2 (Concentration inequality [1]): Assume that h satisfies (B1) and (B2). Let $D = 2\sup_{u\in\mathcal{U}} \|u\|_2$, $C = 32$, $u_0 \in \mathcal{U}$, and Σ_0 denote the covariance matrix of $h(u_0, \cdot)$, and define the constant

$$\tilde{C} = \left(8CD^{\frac{\alpha+1}{2}} d^{\frac{1}{2}} \mathbb{E}\left[M^2\right]^{\frac{1}{2}} \alpha^{-\frac{1}{2}} + \text{Trace}\left(\Sigma_0\right)^{\frac{1}{2}}\right). \tag{4}$$

Let $M \in \mathbb{N}$ be the number of iid samples ω^i and define

$$\epsilon_M = 2\tilde{C}M^{-\frac{1}{4}}, \quad \delta_M = \exp\left(-\frac{\tilde{C}^2}{2\bar{h}^2}M^{\frac{1}{4}}\right). \tag{5}$$

Then, with probability at least $1-\delta_M$ over the M iid samples ω^i,

$$\sup_{u\in\mathcal{U}}\left|\frac{1}{M}\sum_{i=1}^M h(u, \omega^i) - \mathbb{E}[h(u, \cdot)]\right| \leq \epsilon_M. \tag{6}$$

Lemma 2 provides a high-probability finite-sample error bound for the sample average approximation of the expected value of $h(u, \cdot)$ that holds uniformly over all $u \in \mathcal{U}$. It corresponds to Proposition 3.6 in [1] with $\epsilon = 1/4$. Note that $\epsilon_M \to 0$ and $\delta_M \to 0$ as $M \to \infty$.

III. Proof of Lemma 1

For conciseness, we use notations from the main paper that we do not reintroduce. As a preliminary step, we define the function $Z : \mathcal{U} \times \Omega \to \mathbb{R}$ by

$$Z_u(\omega) = \sup_{s\in[0,T]} G(x_u(s, \omega), \xi(\omega)), \tag{7}$$

which is clearly bounded and Hölder continuous under assumptions **(A1-A4)** (see the main paper and the proof of Theorem 1). Then, given a large-enough compact set $\mathcal{T} \subset \mathbb{R}$, we define the measurable function $g : (\mathcal{T} \times \mathcal{U}) \times \Omega \mapsto \mathbb{R}$,

$$g((t, u), \omega) \mapsto t + \alpha^{-1}\max(Z_u(\omega) - t, 0). \tag{8}$$

Note that, since g is a composition of α-Hölder-continuous functions (note that $\max$ is Lipschitz) and $\mathcal{T}$ is bounded, g is also bounded and α-Hölder continuous, i.e., g satisfies assumptions (B1) and (B2) over $(t, u) \in \mathcal{T} \times \mathcal{U}$. We thus have the following result.

Corollary 1 (Finite-sample error bound): Given M iid samples $\omega^i \in \Omega$, formulate the sample average approximation $\textbf{SOCP}_M$ of **OCP**. Define the map g as in (8). Then, under assumptions **(A1-A4)** (see the main paper), for some finite constants $(\tilde{C}, \bar{h})$ large-enough, for (ϵ_M, δ_M) defined in (5), with probability at least $1-\delta_M$ over the M samples,

$$\sup_{(t,u)\in\mathcal{T}\times\mathcal{U}} \left| \frac{1}{M}\sum_{i=1}^{M} g((t,u),\omega^i) - \mathbb{E}[g((t,u),\cdot)] \right| \leq \epsilon_M. \quad (9)$$

Proof: Under assumptions **(A1-A4)**, the function g defined above satisfies Assumptions (B1) and (B2). The conclusion follows from Lemma 2. ■

The desired result (Lemma 2) follows directly from Corollary 1, i.e., we obtain high-probability finite-sample error bounds for the risk constraints given a solution of the sample-based approximation $\textbf{SOCP}_M(\bar{\omega})$ of **OCP**.

Proof of Lemma 2: For any $u \in \mathcal{U}$, define the random variable $Z_u(\omega) = (7)$ (as in the paper), so that

$$\text{AV@R}_\alpha\left(\sup_{s\in[0,T]} G(x_{u(\bar{\omega})}(s),\xi)\right) = \text{AV@R}_\alpha\left(Z_{u(\bar{\omega})}\right).$$

Then, define the function $g : (\mathcal{T}\times\mathcal{U})\times\Omega \mapsto \mathbb{R}$ as in (8). Denote the solution to $\textbf{SOCP}_M(\bar{\omega})$ by $(t(\bar{\omega}), u(\bar{\omega}))$, and note that

$$g((t(\bar{\omega}),u(\bar{\omega})),\omega) = t(\bar{\omega}) + \frac{1}{\alpha}\max(Z_{u(\bar{\omega})}(\omega) - t(\bar{\omega}), 0).$$

By the feasibility of $\textbf{SOCP}_M(\bar{\omega})$, we have that

$$\begin{aligned} &\frac{1}{M}\sum_{i=1}^{M} g((t(\bar{\omega}),u(\bar{\omega})),\omega^i) \\ &\quad = t(\bar{\omega}) + \frac{1}{\alpha M}\sum_{i=1}^{M}\max(Z_{u(\bar{\omega})}(\omega^i) - t(\bar{\omega}), 0) \\ &\quad \leq 0. \end{aligned}$$

Moreover, by Corollary (1), with probability at least $1-\delta_M$ over the M samples ω^i,

$$\sup_{(t,u)\in\mathcal{T}\times\mathcal{U}} \left| \frac{1}{M}\sum_{i=1}^{M} g((t,u),\omega^i) - \mathbb{E}[g((t,u),\cdot)] \right| \leq \epsilon_M.$$

By combining the last two inequalities, we conclude that

$$\mathbb{E}[g((t(\bar{\omega}),u(\bar{\omega})),\cdot)] \leq \epsilon_M$$

with probability at least $1-\delta_M$.

Finally, for $\mathcal{T}$ large-enough (this bounded set $\mathcal{T}$ exists as discussed in the proof of Theorem 1 in the main paper, see also [2]), by definition,

$$\begin{aligned} \text{AV@R}_\alpha(Z_{u(\bar{\omega})}) &= \inf_{t\in\mathbb{R}}\left(t + \frac{1}{\alpha}\mathbb{E}[\max(Z_{u(\bar{\omega})} - t, 0)]\right) \\ &= \inf_{t\in\mathcal{T}}\left(t + \frac{1}{\alpha}\mathbb{E}[\max(Z_{u(\bar{\omega})} - t, 0)]\right) \\ &\leq t(\bar{\omega}) + \frac{1}{\alpha}\mathbb{E}[\max(Z_{u(\bar{\omega})} - t(\bar{\omega}), 0)]. \\ &= \mathbb{E}[t(\bar{\omega}) + \alpha^{-1}\max(Z_{u(\bar{\omega})} - t(\bar{\omega}), 0)]. \\ &= \mathbb{E}[g((t(\bar{\omega}),u(\bar{\omega})),\cdot)] \\ &\leq \epsilon_M \end{aligned}$$

with probability at least $1-\delta_M$, concluding the proof. ■

References

[1] T. Lew, R. Bonalli, and M. Pavone. Sample average approximation for stochastic programming with equality constraints. Available at `https://arxiv.org/abs/2206.09963`, 2022.

[2] A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on Stochastic Programming: Modeling and Theory, Second Edition*. Society for Industrial and Applied Mathematics, 2014.