Resolvent Ellipsoid for Minty Set Inclusions

Ashkan Soleymani **Gabriele Farina** **Patrick Jaillet**
Massachusetts Institute of Technology
{ashkanso,gfarina,jaillet}@mit.edu

Abstract

We study Minty set inclusion on a compact convex set $\mathcal{X} \subseteq \mathbb{R}^d$ with a set-valued operator $A : \mathcal{X} \rightrightarrows \mathbb{R}^d$. Under the Minty promise, the goal is to compute $x \in \mathcal{X}$ and $u \in A(x)$ such that $\sup_{y \in \mathcal{X}} \langle u, x - y \rangle \leq \varepsilon$. Our approach is geometric and we work in the fixed-scale constrained-resolvent oracle model, where a query at anchor a is assumed to return a resolvent point x^+ together with the Yosida vector $g_\eta(a) = (a - x^+)/\eta$. We prove that each oracle response yields a sharp dichotomy: either the returned point provides an ε-SVI certificate, or the Yosida vector defines a strict separating normal for the Minty set. Thus, a single black-box resolvent query supplies either a certified candidate solution or a valid geometric cut. Combining this separation principle with the ellipsoid method, we obtain RESOLVENT-ELLIPSOID, an algorithm that computes an ε-SVI in $\mathcal{O}(\mathrm{poly}(d, \log(1/\varepsilon)))$ black-box oracle calls to the fixed-scale constrained resolvent. For monotone inclusions, our algorithm gives, to our knowledge, the first $\mathcal{O}(\mathrm{poly}(d, \log(1/\varepsilon)))$ black-box guarantee in the fixed-scale resolvent model.

1 Introduction

Black-box resolvent access is a classical oracle model for monotone inclusion problems, dating back to the proximal point method of Rockafellar [1976]. In this model, the basic oracle operation is to solve a regularized inclusion (resolvent subproblem), whereas the standard first-order oracle for variational inequalities evaluates a single-valued operator at a given point. For classical proximal-point methods, the known worst-case convergence rates under monotonicity alone are sublinear. After k resolvent evaluations, the proximal point method has a tight $\mathcal{O}(1/k)$ rate for the squared fixed-point residual [Gu and Yang, 2020], while accelerated variants achieve $\mathcal{O}(1/k^2)$ [Kim, 2021, Park and Ryu, 2022]. These rates translate into $\mathcal{O}(1/\varepsilon)$ and $\mathcal{O}(1/\sqrt{\varepsilon})$ resolvent evaluations, respectively, to obtain an iterate whose squared fixed-point residual is at most ε. Moreover, Park and Ryu [2022] establish matching complexity bounds for general nonexpansive fixed-point schemes. Linear convergence, which yields logarithmic dependence on $1/\varepsilon$, is typically established under additional regularity assumptions, such as strong monotonicity, contractivity, metric subregularity, or Hölder-type growth conditions [Rockafellar, 1976, Leventhal, 2009, Kim, 2021, Park and Ryu, 2022].

For variational inequalities (VIs) with a single-valued operator $F : \mathcal{X} \to \mathbb{R}^d$, the standard first-order oracle returns $F(x)$ at a queried point x. For Lipschitz monotone operators, Halpern-type methods achieve near-optimal $\widetilde{\mathcal{O}}(1/\varepsilon)$ complexity in terms of an operator residual [Diakonikolas, 2020]. More recently, Anagnostides et al. [2025] obtained the first $\mathrm{poly}(d, \log(1/\varepsilon))$ point-evaluation algorithm for Lipschitz VIs under the Minty condition. Their algorithm computes an ε-approximate solution of the Stampacchia variational inequality (ε-SVI), namely a point $x \in \mathcal{X}$ satisfying $\sup_{y \in \mathcal{X}} \langle F(x), x - y \rangle \leq \varepsilon$. The main geometric challenge is to construct separating cuts with positive margin. A cut through an ellipsoid center a_t with normal $F(a_t)$ may pass through a Minty solution and therefore fail to preserve any positive-radius neighborhood of the Minty set. The extra-gradient ellipsoid method addresses this issue by evaluating F at a corrected point $\tilde{a}_t$ and using $F(\tilde{a}_t)$ as the cutting normal.

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

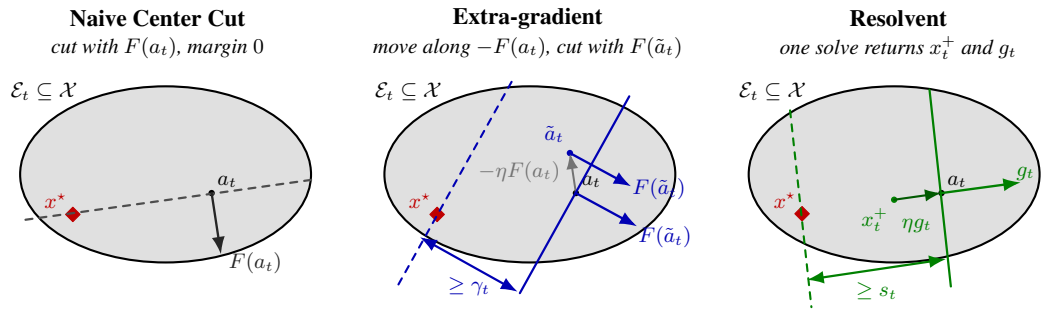

Figure 1: How extra-gradient and resolvent repair the same zero-margin obstruction. Left: the naive cut through a_t with normal $F(a_t)$ can pass through a Minty solution $x^\star$, so it separates but does not protect any positive-radius neighborhood of the Minty set. Middle: extra-gradient uses the same center direction $F(a_t)$ as the naive panel, but first moves from a_t to $\tilde{a}_t$ along $-F(a_t)$ and then cuts through a_t with the corrected normal $F(\tilde{a}_t)$; the dashed parallel support line through $x^\star$ makes the margin explicit, and γ_t is the perpendicular offset between that support line and the strict cut. Right: the constrained resolvent returns $x_t^+ = J_{\eta(F+N_{\mathcal{X}})}(a_t)$ and the Yosida vector $g_t = (a_t - x_t^+)/\eta$ in one shot; hence the correction segment $x_t^+ \to a_t$ and the normal g_t are collinear, the dashed parallel support line through $x^\star$ shows the cut margin $s_t = \|a_t - x_t^+\| = \eta\|g_t\|$. Extra-gradient and resolvent therefore repair the same geometric failure (lack of margin) by explicit versus implicit correction.

For structured inclusion problems arising in composite optimization, primal-dual saddle-point systems, and KKT formulations, computation often relies on regularized subproblem solves, such as proximal maps, splitting steps, or conic routines for strongly regularized instances. Can the polylogarithmic dependence on $1/\varepsilon$ achieved for single-valued Minty variational inequalities also be obtained for set-valued Minty inclusions using only fixed-scale resolvent queries?

We answer this question affirmatively. We consider a set-valued operator $A : \mathcal{X} \rightrightarrows \mathbb{R}^d$ on a compact convex set $\mathcal{X}$ and study the following Minty inclusion problem: under the promise that the Minty set is nonempty, compute a point $x \in \mathcal{X}$ and a vector $u \in A(x)$ such that

$$\sup_{y \in \mathcal{X}} \langle u, x - y \rangle \leq \varepsilon.$$

Our main result gives a $\mathrm{poly}(d, \log(1/\varepsilon))$ algorithm in the fixed-scale resolvent model.

The key observation is that a single regularized solve simultaneously produces both an approximate solution and a separating normal. Specifically, if

$$a = x^+ + \eta(u^+ + n^+), \qquad u^+ \in A(x^+), \qquad n^+ \in N_{\mathcal{X}}(x^+),$$

then the Yosida vector

$$g_\eta(a) := u^+ + n^+$$

provides the candidate point x^+ together with the normal that defines the cut. In this sense, one resolvent query automatically performs the correction that extra-gradient methods construct explicitly. Figure 1 illustrates this shared margin geometry. In the left panel, the naive cut through a_t with normal $F(a_t)$ can pass through a Minty solution $x^\star$, yielding zero margin. In the middle panel, the extra-gradient method moves from a_t to $\tilde{a}_t$ along the direction $F(a_t)$ and obtains a strict cut whose margin γ_t is the perpendicular distance from $x^\star$ to the cut. In the right panel, the resolvent analogue produces a correction segment $x_t^+ \to a_t$ that is collinear with the normal g_t, and the cut margin is the offset s_t. Thus, the resolvent oracle directly provides the information needed to construct a positive-margin cut, without requiring a separate extra-gradient correction.

The geometry is clearest in the unconstrained gradient case $F = \nabla f$. At a center a_t, the naive ellipsoid cut with normal $\nabla f(a_t)$ may have zero margin: it can separate but still pass through a Minty solution. Extra-gradient repairs this by an *explicit predictor-corrector* step. It first moves *forward* to $\widetilde{a}_t = a_t - \eta \nabla f(a_t)$, and then cuts through a_t using the corrected normal $\nabla f(\widetilde{a}_t)$. The Lipschitz assumption is precisely what makes this forward correction controlled and converts separation into

separation with positive margin. The resolvent step performs the analogous correction *implicitly*. Instead of evaluating the field at a_t, it computes x_t^+ from the backward equation

$$x_t^+ = a_t - \eta \nabla f(x_t^+), \qquad g_t := \frac{a_t - x_t^+}{\eta} = \nabla f(x_t^+).$$

Equivalently, when f is convex, $x_t^+ \in \arg\min_x \left\{ f(x) + \frac{1}{2\eta}\|x - a_t\|^2 \right\}$. Thus the corrected point and the corrected normal are produced in one step: the displacement $a_t - x_t^+ = \eta g_t$ is already collinear with the cutting normal. Consequently, the strict margin is $s_t = \|a_t - x_t^+\| = \eta\|g_t\|$. This is the key distinction. Extra-gradient obtains margin by an *explicit forward correction*, justified by Lipschitz continuity; the resolvent obtains the same geometric effect from an *implicit backward correction* encoded by the proximal solve. This backward formulation is what remains meaningful in the set-valued setting, where there may be no single vector $F(a_t)$ to evaluate and no Lipschitz continuity assumption to invoke.

We focus on the following set-valued problem.

Definition 1 (Minty set inclusion)**.** *Let $\mathcal{X} \subseteq \mathbb{R}^d$ be compact and convex, and let $A : \mathcal{X} \rightrightarrows \mathbb{R}^d$ have nonempty values. The goal is to compute, under the promise $\mathcal{K} \neq \varnothing$, a point $x \in \mathcal{X}$ and a witness $u \in A(x)$ such that*

$$\sup_{y \in \mathcal{X}} \langle u, x - y \rangle \leq \varepsilon.$$

We call such a pair an ε-Stampacchia solution *or* ε-SVI*.*

When $A(x) = \{F(x)\}$ is single-valued, this is the usual Minty-VI task on $\mathcal{X}$. The set-valued formulation also captures subdifferentials, composite operators, and KKT maps, which is precisely the setting in which resolvent access is natural [Minty, 1962, Rockafellar, 1976]. The contribution is therefore complementary to the point-evaluation model of Anagnostides et al. [2025] as we show that fixed-scale regularized solves become a sufficient primitive for Minty-SVI computation in the broader set-valued setting.

Our Contributions. We demonstrate on a compact convex domain $\mathcal{X}$, *fixed-scale constrained resolvent access together with linear optimization over $\mathcal{X}$* yields an ε-SVI in $\mathrm{poly}(d, \log(1/\varepsilon))$ time under the Minty promise. For monotone A, the Minty set coincides with the constrained zero set

$$K \;=\; \{x \in \mathcal{X} : 0 \in A(x) + N_{\mathcal{X}}(x)\},$$

so the framework specializes directly to feasible compact constrained monotone inclusions of the form $0 \in A(x) + N_{\mathcal{X}}(x)$. Thus, to the best of our knowledge, the monotone specialization of our framework gives the *first* $\mathrm{poly}(d, \log(1/\varepsilon))$ black-box guarantee in the fixed-scale constrained resolvent model for feasible compact constrained monotone inclusions, without assuming Lipschitz point evaluation, strong monotonicity, cocoercivity, metric subregularity, or any other error-bound or restrictive geometry conditions.

Our other contributions are as follows.

- We show that a single constrained resolvent query provides both an acceptance certificate and a strict separator for the Minty set. This is the geometric core of the paper.
- We obtain exact and inexact outer theorems with polylogarithmic accuracy dependence. In the nontrivial regime $\varepsilon \leq R^2/\eta$, the outer complexity is $\mathcal{O}\left(d^2 \log \frac{R^2}{\eta\varepsilon}\right)$, and any $\mathrm{poly}(d, \log(1/\delta))$ routine for approximate constrained resolvents lifts to a $\mathrm{poly}(d, \log(1/\varepsilon))$ algorithm for Minty-SVI. In particular, this brings polylogarithmic-accuracy guarantees into the native resolvent-oracle literature for compact constrained monotone inclusions.
- We extend the geometry to a set-valued weak-Minty condition and prove an either-ε-SVI-or-certificate alternative when the Minty condition fails. As a consequence, we broaden the regime of the weak-Minty condition for VIs with polylogarithmic-accuracy tractability (Remark 14).
- We prove a converse on ρ-hypomonotone classes and briefly record canonical oracle realizations.

The appendix develops the cutting-plane upgrade and exact certificate duality, and presents representative realizations of the resolvent oracle. It also provides a counterexample to a full converse under the Minty assumption alone and establishes a sharp linear convergence rate under a Yosida error bound.

2 Related Work

Proximal and resolvent methods have a long history in monotone operator theory, from Moreau's proximal mapping and Rockafellar's proximal point method to splitting and primal–dual schemes such as Douglas–Rachford and Chambolle–Pock [Moreau, 1965, Rockafellar, 1976, Lions and Mercier, 1979, Eckstein and Bertsekas, 1992, Chambolle and Pock, 2011, Bauschke and Combettes, 2017]. These methods rely on regularized subproblem solves, either through the resolvent of the full operator or through proximal evaluations of individual components. Under monotonicity alone, the known worst-case residual convergence rates for classical proximal-point and fixed-point iterations are algebraic, yielding oracle complexity polynomial in $1/\epsilon$ for the corresponding residual criteria [Gu and Yang, 2020, Kim, 2021, Park and Ryu, 2022].

Cutting-plane methods provide another approach to variational inequalities with single-valued operators, using point evaluations of F to construct separating hyperplanes. Lüthi [1985] applied the ellipsoid method to monotone VIs, although the logarithmic-accuracy ϵ-SVI guarantee in that analysis requires strong monotonicity, as noted by Anagnostides et al. [2025]. Subsequent analytic-center methods considered broader classes of VIs. Goffin et al. [1997] studied continuous pseudomonotone VIs and obtained nonasymptotic guarantees under the stronger assumption of pseudo-cocoercivity. Nesterov and Vial [1999] developed homogeneous analytic-center cutting-plane methods with efficiency estimates for variational inequalities, feasibility problems, and constrained minimization. Shen and Pang [2012] considered continuous operators with approximate point evaluations and established convergence of a proximal analytic-center cutting-plane method. The closest result to ours is Anagnostides et al. [2025], who obtained the first $\mathrm{poly}(d, \log(1/\epsilon))$ algorithm for ϵ-SVIs of single-valued Lipschitz operators under the Minty condition using point-evaluation access. Our work establishes the corresponding polylogarithmic accuracy dependence for set-valued Minty inclusions using fixed-scale resolvent queries.

3 Notation and Problem Setup

Notation. We write $\langle \cdot, \cdot \rangle$ for the Euclidean inner product and $\|\cdot\|$ for the Euclidean norm. For $x \in \mathbb{R}^d$ and $r > 0$,

$$B_r(x) := \left\{ z \in \mathbb{R}^d : \|z - x\| \le r \right\}$$

denotes the closed Euclidean ball of radius r centered at x. In particular, $B_R(0)$ is the radius-R ball centered at the origin. For sets $C, D \subseteq \mathbb{R}^d$, we write $C + D := \{c + d : c \in C,\ d \in D\}$. The graph of a set-valued operator is $\operatorname{gra} A := \left\{ (x, u) \in \mathcal{X} \times \mathbb{R}^d : u \in A(x) \right\}$.

We assume throughout that

$$B_1(0) \subseteq \mathcal{X} \subseteq B_R(0), \qquad R \ge 1. \tag{1}$$

As usual, this normalization can be enforced by a standard affine preprocessing for well-bounded convex sets [Grötschel et al., 1993]. For $x \in \mathcal{X}$, the normal cone is

$$N_{\mathcal{X}}(x) := \left\{ n \in \mathbb{R}^d : \langle n, y - x \rangle \le 0 \text{ for all } y \in \mathcal{X} \right\}.$$

Definition 2 (Constrained resolvent and Yosida vector)**.** *Set $T := A + N_{\mathcal{X}}$. For a scale $\eta > 0$ and anchor $a \in \mathbb{R}^d$, the constrained resolvent set is*

$$J_\eta(a) := (I + \eta T)^{-1}(a) = \{ x \in \mathcal{X} : \exists u \in A(x),\ n \in N_{\mathcal{X}}(x) \text{ s.t. } a = x + \eta(u + n) \}.$$

Any choice of $x^+ \in J_\eta(a)$ together with $u^+ \in A(x^+)$ and $n^+ \in N_{\mathcal{X}}(x^+)$ satisfying $a = x^+ + \eta(u^+ + n^+)$ induces the Yosida vector

$$g_\eta(a) := u^+ + n^+ = \frac{a - x^+}{\eta}.$$

When T is maximal monotone, J_η is the classical single-valued resolvent and g_η is the corresponding Yosida approximation [Moreau, 1965, Rockafellar, 1976, Bauschke and Combettes, 2017]. Under the Minty assumption alone, the same formulas apply, but $J_\eta(a)$ may contain multiple points. The oracle selects a solution $x^+ \in J_\eta(a)$, which determines the Yosida vector $g_\eta(a)$.

Definition 3 (Minty set). *The Minty set of A on $\mathcal{X}$ is*

$$\mathcal{K} := \{x \in \mathcal{X} : \langle v, y - x\rangle \geq 0 \text{ for all } y \in \mathcal{X},\ v \in A(y)\}.$$

The Minty set is always compact and convex because it is an intersection of $\mathcal{X}$ with closed halfspaces indexed by graph points of A; the short proof is deferred to Appendix A.

We work in the fixed-scale model. Exact resolvent access at arbitrarily large scales makes the approximation problem essentially one-shot; Proposition 21 records this observation. We therefore study fixed-scale access.

Assumption 4 (Exact resolvent oracle). *Fix $\eta > 0$. For every rational anchor $a \in \mathbb{R}^d$, an oracle returns a triple (x^+, u^+, n^+) satisfying*

$$x^+ \in \mathcal{X}, \qquad u^+ \in A(x^+), \qquad n^+ \in N_{\mathcal{X}}(x^+), \qquad a = x^+ + \eta(u^+ + n^+).$$

We write $x^+ = J_\eta(a)$ for the returned selection and

$$g_\eta(a) := u^+ + n^+ = \frac{a - x^+}{\eta}.$$

We also assume access to linear optimization over $\mathcal{X}$.

Black-box resolvent oracle access is a standard algorithmic primitive in the study of monotone inclusion problems Minty [1962], Rockafellar [1976], Luque [1984], Qin et al. [2010], Briceno-Arias and Combettes [2011], Boţ and Csetnek [2014], Gu and Yang [2020], Bui and Combettes [2020], Kim [2021], Bauschke et al. [2021], Park and Ryu [2022], Sadiev et al. [2024], Boţ and Nguyen [2026]. For the definition of the δ-approximate resolvent oracle, we refer to Assumption 22 in Appendix A.

4 Resolvent Geometry

Fix $\eta > 0$ and an exact resolvent selection as in Assumption 4. For an anchor $a \in \mathbb{R}^d$, let

$$x^+ := J_\eta(a), \qquad g_\eta(a) := \frac{a - x^+}{\eta}.$$

Definition 5 (Yosida gap). *For an anchor $a \in \mathbb{R}^d$, define*

$$\mathrm{Gap}_\eta(a) := \max_{y \in \mathcal{X}} \langle g_\eta(a), a - y\rangle.$$

The first lemma is the acceptance test that serves as a Gap certificate.

Lemma 6. *If $\mathrm{Gap}_\eta(a) \leq \varepsilon$, then the resolvent point $x^+ = J_\eta(a)$ is an ε-SVI point.*

Proof. Write $g_\eta(a) = u^+ + n^+$ with $u^+ \in A(x^+)$ and $n^+ \in N_{\mathcal{X}}(x^+)$. For any $y \in \mathcal{X}$,

$$\langle n^+, y - x^+\rangle \leq 0 \qquad \Longrightarrow \qquad \langle u^+, x^+ - y\rangle \leq \langle u^+ + n^+, x^+ - y\rangle.$$

Using $a = x^+ + \eta g_\eta(a)$,

$$\langle u^+, x^+ - y\rangle \leq \langle g_\eta(a), a - y\rangle - \langle g_\eta(a), a - x^+\rangle = \langle g_\eta(a), a - y\rangle - \eta \|g_\eta(a)\|^2 \leq \mathrm{Gap}_\eta(a).$$

Thus $\langle u^+, x^+ - y\rangle \leq \varepsilon$ for all $y \in \mathcal{X}$. □

The next lemma is the key geometric observation that the constrained resolvent itself gives a strict separating direction for the Minty set.

Lemma 7 (Exact strict separation). *Let $a \in \mathbb{R}^d$ and let $x^+ = J_\eta(a)$. For every $x \in \mathcal{K}$,*

$$\langle g_\eta(a), a - x\rangle \geq \eta \|g_\eta(a)\|^2.$$

Equivalently, if $c(a) := \frac{g_\eta(a)}{\|g_\eta(a)\|}$, then

$$\langle c(a), x - a\rangle \leq -\eta \|g_\eta(a)\|.$$

Proof. Write $g_\eta(a) = u^+ + n^+$ with $u^+ \in A(x^+)$ and $n^+ \in N_{\mathcal{X}}(x^+)$. Since $x \in \mathcal{K}$ and $u^+ \in A(x^+)$,

$$\langle u^+, x^+ - x \rangle \geq 0.$$

Since $n^+ \in N_{\mathcal{X}}(x^+)$ and $x \in \mathcal{X}$,

$$\langle n^+, x^+ - x \rangle \geq 0.$$

Adding the two inequalities gives

$$\langle g_\eta(a), x^+ - x \rangle \geq 0.$$

Therefore

$$\langle g_\eta(a), a - x \rangle = \langle g_\eta(a), a - x^+ \rangle + \langle g_\eta(a), x^+ - x \rangle \geq \langle g_\eta(a), a - x^+ \rangle = \eta \, \|g_\eta(a)\|^2 .$$

The normalized form follows by dividing by $\|g_\eta(a)\|$. □

To turn a failed acceptance test into a uniform central cut, define

$$s(a) := \|a - J_\eta(a)\| = \eta \, \|g_\eta(a)\| ,$$

which allows us to derive a quantitative margin from a failed gap test as follows.

Lemma 8. *If* $\mathrm{Gap}_\eta(a) > \varepsilon$*, then*

$$s(a) > \rho_\eta(\varepsilon, R) := \sqrt{R^2 + \eta\varepsilon} - R.$$

Consequently, the halfspace

$$H(a) := \{ z \in \mathbb{R}^d : \langle g_\eta(a), z - a \rangle \leq 0 \}$$

contains the thickened body $\mathcal{K} + B_{\rho_\eta(\varepsilon,R)}(0)$.

Proof sketch. The upper bound

$$\mathrm{Gap}_\eta(a) \leq \frac{1}{\eta}\Big(s(a)^2 + 2Rs(a) \Big)$$

follows from $\|x^+ - y\| \leq 2R$ for $x^+, y \in \mathcal{X}$. Thus $\mathrm{Gap}_\eta(a) > \varepsilon$ implies $s(a)^2 + 2Rs(a) - \eta\varepsilon > 0$, hence $s(a) > \sqrt{R^2 + \eta\varepsilon} - R$. Now take $z = x + h$ with $x \in \mathcal{K}$ and $\|h\| \leq \rho_\eta(\varepsilon, R)$. By Lemma 7,

$$\langle g_\eta(a), z - a \rangle \leq -\eta \, \|g_\eta(a)\|^2 + \rho_\eta(\varepsilon, R) \, \|g_\eta(a)\| < 0$$

because $\eta \, \|g_\eta(a)\| = s(a) > \rho_\eta(\varepsilon, R)$. The detailed proof is in Appendix B. □

The geometry relies only on Minty feasibility and the normal-cone relation. No extra-gradient step and no Lipschitz continuity of A are used in Lemmas 6 to 8. This is because the resolvent simultaneously provides both a gap certificate (Lemma 6) and a strict separation (Lemmas 7 and 8).

5 Exact and Inexact Outer Theorems

The outer method is now simple. At the current ellipsoid center a_t, query the (exact or approximate) constrained resolvent, compute the Yosida gap by linear optimization over $\mathcal{X}$, and either accept or insert the central cut through a_t given by the returned Yosida vector.

Theorem 9 (Exact resolvent ellipsoid)**.** *Under Assumption 4 and the Minty condition* $\mathcal{K} \neq \varnothing$*, the exact outer method returns an* ε*-SVI after* $\mathcal{O}\Big(d^2 \log \frac{R + \rho_\eta(\varepsilon,R)}{\rho_\eta(\varepsilon,R)} \Big)$ *iterations. Under the nontrivial regime* $\varepsilon \leq R^2/\eta$*, this simplifies to*

$$\mathcal{O}\left(d^2 \log \frac{R^2}{\eta\varepsilon} \right).$$

Each iteration uses one exact resolvent call and one linear-optimization call.

Proof sketch. Lemma 8 shows that every rejected iterate generates a central cut containing $\mathcal{K} + B_{\rho_\eta(\varepsilon,R)}(0)$. Hence every maintained ellipsoid contains a fixed full-dimensional convex body with inscribed radius $\rho_\eta(\varepsilon, R)$ and outer radius at most $R + \rho_\eta(\varepsilon, R)$. The standard central-cut ellipsoid volume argument [Grötschel et al., 1993] therefore forces termination after the stated number of iterations. When the algorithm terminates, Lemma 6 upgrades the final accepted iterate to an ε-SVI. Full details are in Appendix B. □

We now move to the approximate oracle model. For an approximate triple $(\hat{x}, \hat{u}, \hat{n})$, define

$$\widehat{\mathrm{Gap}}_\eta(a) := \max_{y \in \mathcal{X}} \langle \hat{g}(a), a - y \rangle, \qquad \hat{g}(a) := \hat{u} + \hat{n}.$$

Theorem 10 (Robust outer theorem)**.** *Assume the Minty condition, the nontrivial regime*

$$\varepsilon \le \frac{R^2}{\eta}, \tag{2}$$

and access to a δ-approximate constrained resolvent oracle at scale η with $\delta \le \frac{\eta\varepsilon}{16R}$. Run the outer method from the initial ball $B_{2R}(0)$, accept whenever $\widehat{\mathrm{Gap}}_\eta(a_t) \le \varepsilon/2$, and otherwise cut with the halfspace $\{z : \langle \hat{g}(a_t), z - a_t \rangle \le 0\}$. Then the method returns an ε-SVI after $\mathcal{O}\Big(d^2 \log \frac{R^2}{\eta\varepsilon}\Big)$ iterations.

Proof sketch. Two robust facts are proved in Appendix B. First, if $\widehat{\mathrm{Gap}}_\eta(a) \le \varepsilon/2$ and $\delta \le \eta\varepsilon/(8R)$, then $\hat{x}$ is already an ε-SVI. Second, if $\widehat{\mathrm{Gap}}_\eta(a) > \varepsilon/2$ and $\delta \le \eta\varepsilon/(16R)$, then the rejected cut still contains $\mathcal{K} + B_{\eta\varepsilon/(16R)}(0)$. Thus the same ellipsoid volume argument as in Theorem 9 applies, now with fixed thickening radius $\eta\varepsilon/(16R)$. □

Corollary 11. *Fix $\eta > 0$. Suppose one can compute a δ-approximate constrained resolvent selection of $A + N_{\mathcal{X}}$ at scale η in time*

$$T_{\mathrm{res}}(d, \log(1/\delta)).$$

Then, under the Minty condition, one can compute an ε-SVI in time

$$\mathcal{O}\Big(d^2 \log \frac{R^2}{\eta\varepsilon}\Big) \cdot \Big(T_{\mathrm{res}}(d, \log(R/(\eta\varepsilon))) + T_{\mathrm{lin}}(d)\Big),$$

where $T_{\mathrm{lin}}(d)$ is the cost of linear optimization over $\mathcal{X}$. In particular, if T_{res} is polynomial in d and logarithmic in $1/\delta$, then the Minty-SVI problem is polynomial in d and logarithmic in $1/\varepsilon$.

Appendix E shows that the same oracle can also be plugged into the cutting-plane method of Jiang et al. [2020], reducing the number of outer oracle calls from $\mathcal{O}(d^2 \log(\cdot))$ to $\mathcal{O}(d \log(\cdot))$.

6 Weak Minty Beyond Point Evaluation

Anagnostides et al. [2025, Section 4.3] also proves a polylogarithmic theorem under a constrained weak-Minty condition for single-valued Lipschitz fields. In the resolvent model there is a natural set-valued extension.

Definition 12 (Graph weak Minty)**.** *For a graph point $(y, v) \in \mathrm{gra}\, A$, define its witness-specific SVI gap by*

$$\mathrm{gap}_A(y, v) := \max_{z \in \mathcal{X}} \langle v, y - z \rangle.$$

For $\kappa \ge 0$, define the weak-Minty set

$$\mathcal{K}_\kappa := \big\{x \in \mathcal{X} : \langle v, y - x \rangle \ge -\kappa\, \mathrm{gap}_A(y, v)^2 \text{ for all } (y, v) \in \mathrm{gra}\, A\big\}.$$

*We say that A satisfies the κ-*graph weak-Minty condition *if $\mathcal{K}_\kappa \neq \varnothing$. In particular, $\mathcal{K}_0 = \mathcal{K}$.*

When $A(x) = \{F(x)\}$ is single-valued, $\mathrm{gap}_A(y, F(y)) = \max_{z \in \mathcal{X}} \langle F(y), y - z \rangle$, so Definition 12 coincides exactly with the constrained weak-Minty condition studied in Anagnostides et al. [2025]. The resolvent model extends that condition to set-valued, proximable, and KKT-style settings.

The key algebraic fact is that one exact resolvent call already returns a graph point (x^+, u^+) whose witness-specific gap is controlled by the Yosida norm,

$$\mathrm{gap}_A(x^+, u^+) \le 2R \, \|g_\eta(a)\| .$$

Hence, the weak-Minty loss is quadratic in exactly the quantity that drives the exact Minty margin, which we use in our analysis.

Theorem 13 (Weak-Minty outer theorem)**.** *Let $\lambda_\kappa := 1 - \frac{4\kappa R^2}{\eta}$. Assume Assumption 4 and the κ-graph weak-Minty condition with $\lambda_\kappa > 0$. Then every rejected anchor a produces a central cut containing $\mathcal{K}_\kappa + B_{\lambda_\kappa \rho_\eta(\varepsilon, R)}(0)$. Consequently, the exact outer method returns an ε-SVI after $\mathcal{O}\Big(d^2 \log \frac{R + \lambda_\kappa \rho_\eta(\varepsilon, R)}{\lambda_\kappa \rho_\eta(\varepsilon, R)}\Big)$ iterations. Under the nontrivial regime $\varepsilon \le R^2/\eta$, this is equivalent to $\mathcal{O}\Big(d^2 \log \frac{R^2}{(\eta - 4\kappa R^2)\varepsilon}\Big)$ iterations.*

Proof sketch. Let $x^+ = J_\eta(a)$ and write $g_\eta(a) = u^+ + n^+$. For any $x \in \mathcal{K}_\kappa$, weak Minty at the graph point (x^+, u^+) and the normal-cone relation give

$$\langle g_\eta(a), a - x \rangle \ge \eta \, \|g_\eta(a)\|^2 - \kappa \, \mathrm{gap}_A(x^+, u^+)^2.$$

Since $\mathrm{gap}_A(x^+, u^+) \le 2R \, \|g_\eta(a)\|$, this becomes

$$\langle g_\eta(a), a - x \rangle \ge (\eta - 4\kappa R^2) \, \|g_\eta(a)\|^2 .$$

Normalizing by $\|g_\eta(a)\|$ yields a strict cut with margin $\lambda_\kappa s(a)$, where $s(a) = \eta \, \|g_\eta(a)\|$. If the gap test rejects, Lemma 8 gives $s(a) > \rho_\eta(\varepsilon, R)$, so the cut contains $\mathcal{K}_\kappa + B_{\lambda_\kappa \rho_\eta(\varepsilon, R)}(0)$. The ellipsoid bound is then identical to Theorem 9. Appendix C gives the full proof and the inexact analogue with admissible graph residual $\delta = \mathcal{O}(\lambda_\kappa \eta \varepsilon / R)$. □

Remark 14. *or single-valued L-Lipschitz F, choosing $\eta \asymp 1/L$ yields an admissible weak-Minty threshold of order $1/(LR^2)$. This matches the order obtained by Anagnostides et al. [2025] in the regime $B \lesssim RL$. Our resolvent-based threshold depends only on the resolvent scale η and the domain radius R, independently of the operator bound B. Moreover, our guarantee extends to set-valued inclusions, including nonsmooth proximable systems and compact KKT operators, whenever fixed-scale resolvent access is available.*

7 Either an ε-SVI or a Certificate of Minty Infeasibility

If the Minty condition fails, the rejected transcript provides a witness of its violation.

Definition 15 (Strict graph-expected VI certificate)**.** *Let $(x_t, u_t)_{t=1}^T$ satisfy $x_t \in \mathcal{X}$ and $u_t \in A(x_t)$. A finitely supported distribution*

$$\mu = \sum_{t=1}^{T} \lambda_t \, \delta_{(x_t, u_t)}, \qquad \lambda \in \Delta_T,$$

is a γ-strict graph-expected variational inequality certificate *(γ-strict GEVI certificate) if*

$$\sum_{t=1}^{T} \lambda_t \, \langle u_t, x - x_t \rangle \ge \gamma \qquad \textit{for all } x \in \mathcal{X}.$$

A strict GEVI immediately certifies non-Minty. In terms, if $x^\star \in \mathcal{K}$, then every graph point (x_t, u_t) satisfies $\langle u_t, x^\star - x_t \rangle \le 0$, so averaging cannot produce a positive γ.

Theorem 16 (Either an ε-SVI or a non-Minty certificate)**.** *Assume the hypotheses of Theorem 10. Then, after $\mathcal{O}\Big(d^2 \log \frac{R^2}{\eta\varepsilon}\Big)$ iterations and polynomial additional computation, the inexact outer method returns either*

(i) an ε-SVI point, or

(ii) a $\widehat{\gamma}_{\mathrm{cert}}$-strict GEVI certificate supported on at most $d+2$ graph points, where $\widehat{\gamma}_{\mathrm{cert}} := \frac{\eta\varepsilon^2}{512R^2}$.

Hence, outcome (ii) *certifies that $\mathcal{K} = \varnothing$.*

Proof sketch. If the outer method accepts, Theorem 10 already gives an ε-SVI. Suppose instead that all iterates are rejected. Then Appendix D shows the following pointwise lower bound: for every $x \in \mathcal{X}$, $\max_{1\le t\le T} \langle \hat{u}_t, x - \hat{x}_t\rangle \ge \widehat{\gamma}_{\mathrm{cert}}$. Equivalently, $\min_{x\in\mathcal{X}} \max_{t\in[T]} \langle \hat{u}_t, x - \hat{x}_t\rangle \ge \widehat{\gamma}_{\mathrm{cert}}$. Now maximize over the simplex of mixing weights and invoke Sion's minimax theorem [Sion, 1958]. This yields a convex combination of transcript graph points that is a $\widehat{\gamma}_{\mathrm{cert}}$-strict GEVI certificate. Appendix F proves the sharper exact-model version, exact minimax duality under compactness of $\mathrm{gra}\, A$, and support compression to at most $d+2$ graph points. □

8 A Converse on Weakly Monotone Classes

The forward reduction holds under the Minty condition alone, but the reverse direction fails in general, as shown by the counterexample in Appendix H. We therefore establish the converse for weakly monotone operators, also known as hypomonotone operators, which admit the regularization needed for the reduction.

Definition 17 (ρ-weakly monotone)**.** *Let $\rho \ge 0$. A set-valued operator $A : \mathcal{X} \rightrightarrows \mathbb{R}^d$ is* ρ-hypomonotone *if*

$$\langle u - v, x - y\rangle \ge -\rho \|x - y\|^2 \qquad \text{for all } x, y \in \mathcal{X},\ u \in A(x),\ v \in A(y).$$

The case $\rho = 0$ is ordinary monotonicity.

Next, we prove that, under a weak monotonicity assumption, finding a constrained resolvent point is computationally equivalent to solving the original Minty inclusion problem, or more precisely, the Minty VI problem in this case.

Theorem 18 (Converse for weakly monotone)**.** *Fix $\rho \ge 0$ and $\eta > 0$ with $\eta^{-1} > \rho$. Assume A is ρ-hypomonotone and that the constrained resolvent set $J_\eta(a)$ is nonempty. Then $J_\eta(a) = \{x^+\}$ is a singleton. Moreover, if $x \in \mathcal{X}$ is an $\varepsilon_{\mathrm{in}}$-SVI point for the regularized operator $A_{a,\eta}(x) := A(x) + \frac{1}{\eta}(x - a)$, then*

$$\|x - J_\eta(a)\| \le \sqrt{\frac{\varepsilon_{\mathrm{in}}}{\eta^{-1} - \rho}}.$$

Consequently, one recovers a δ-accurate constrained resolvent point by solving the regularized weakly monotone instance to accuracy $\varepsilon_{\mathrm{in}} = (\eta^{-1} - \rho)\delta^2$.

Proof sketch. The regularized operator $A_{a,\eta} + N_{\mathcal{X}}$ is $(\eta^{-1} - \rho)$-strongly monotone because A is ρ-hypomonotone and $N_{\mathcal{X}}$ is monotone; this gives uniqueness of $J_\eta(a)$. Now compare an $\varepsilon_{\mathrm{in}}$-SVI point x for $A_{a,\eta}$ against the exact resolvent point x^+. The SVI inequality at x and the resolvent KKT relation at x^+ imply

$$\varepsilon_{\mathrm{in}} \ge \left\langle u - u^+, x - x^+\right\rangle + \frac{1}{\eta}\left\|x - x^+\right\|^2,$$

with $u \in A(x)$ and $u^+ \in A(x^+)$. Hypomonotonicity bounds the first term below by $-\rho\|x - x^+\|^2$, yielding

$$\varepsilon_{\mathrm{in}} \ge (\eta^{-1} - \rho)\left\|x - x^+\right\|^2.$$

The full proof is in Appendix B. □

Together, the forward and reverse reductions establish a computational equivalence between solving the Minty VI problem and computing constrained resolvent points for weakly monotone operators, which also include nonmonotone instances.

9 Oracle Realizations

We present representative settings in which the fixed-scale constrained resolvent oracle can be implemented by solving the anchored subproblem to the required graph-residual accuracy, with running time polylogarithmic in the inverse accuracy.

Theorem 19 (Conic-representable regularized subproblems)**.** *Suppose that, for a fixed scale $\eta > 0$, each anchored resolvent subproblem for $A + N_{\mathcal{X}}$ admits a polynomial-size LP, SOCP, SDP, or more generally self-concordant conic representation whose data depend affinely on the anchor. Then standard interior-point methods compute a graph-residual δ-approximate constrained resolvent in time* $\mathrm{poly}(\mathrm{size}(\mathrm{input}), d, \log(1/\delta))$ *[Nesterov and Nemirovskii, 1994]. Consequently, under the Minty condition, the associated Minty-SVI problem is solvable in time* $\mathrm{poly}(\mathrm{size}(\mathrm{input}), d, \log(1/\varepsilon))$.

Appendix G presents fixed-scale resolvent realizations for feasibility problems involving normal cones, composite convex subdifferential inclusions, and compact primal–dual KKT systems. For the single-valued Lipschitz VI families studied by Anagnostides et al. [2025], Section G.2 establishes efficient fixed-scale resolvent computation, and Corollary 45 shows how to obtain the graph-residual oracle required by Corollary 11 on projection-friendly compact domains.

10 Conclusion

We established that fixed-scale resolvent access suffices to compute approximate Stampacchia solutions for set-valued inclusions under the Minty condition, without requiring monotonicity. Our algorithm achieves $\mathrm{poly}(d, \log(1/\varepsilon))$ oracle complexity by constructing positive-margin ellipsoid cuts directly from resolvent evaluations. This extends the localization approach of Anagnostides et al. [2025] beyond single-valued Lipschitz VIs to set-valued inclusions, including normal-cone and KKT formulations. An important open question is whether similar strict-cut guarantees can be established under weaker conditions than Minty, and more generally, what assumptions are necessary for $\mathrm{poly}(d, \log(1/\varepsilon))$ query complexity.

Acknowledgement

A.S. and P.J. were partially supported by the ONR grant N00014-24-1-2470. Part of this research was performed while A.S. was visiting the Institute for Mathematical and Statistical Innovation (IMSI), which is supported by the National Science Foundation (Grant No. DMS-2425650). G.F. was supported in part by the National Science Foundation awards CCF-2443068 and IIS-2552046, the Office of Naval Research grant N000142512296, and an AI2050 Early Career Fellowship.

References

Ioannis Anagnostides, Gabriele Farina, Tuomas Sandholm, and Brian Hu Zhang. A polynomial-time algorithm for variational inequalities under the minty condition. *arXiv preprint arXiv:2504.03432*, 2025.

Heinz H Bauschke and Patrick L Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, 2017.

Heinz H Bauschke, Walaa M Moursi, and Xianfu Wang. Generalized monotone operators and their averaged resolvents. *Mathematical Programming*, 189(1):55–74, 2021.

Radu Ioan Boţ and Ernö Robert Csetnek. Forward-backward and tseng's type penalty schemes for monotone inclusion problems. *Set-Valued and Variational Analysis*, 22(2):313–331, 2014.

Radu Ioan Boţ and Dang-Khoa Nguyen. Tikhonov regularization of monotone operator flows not only ensures strong convergence of the trajectories but also speeds up the vanishing of the residuals. *Journal of Complexity*, 94:102029, 2026.

Luis M Briceno-Arias and Patrick L Combettes. A monotone+ skew splitting model for composite monotone inclusions in duality. *SIAM Journal on Optimization*, 21(4):1230–1250, 2011.

Minh N Bui and Patrick L Combettes. Warped proximal iterations for monotone inclusions. *Journal of Mathematical Analysis and Applications*, 491(1):124315, 2020.

Antonin Chambolle and Thomas Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of mathematical imaging and vision*, 40(1):120–145, 2011.

Jelena Diakonikolas. Halpern iteration for near-optimal and parameter-free monotone inclusion and strong solutions to variational inequalities. In *Conference on Learning Theory*, pages 1428–1451. PMLR, 2020.

Jonathan Eckstein and Dimitri P Bertsekas. On the douglas—rachford splitting method and the proximal point algorithm for maximal monotone operators. *Mathematical programming*, 55(1): 293–318, 1992.

Jean-Louis Goffin, Patrice Marcotte, and Daoli Zhu. An analytic center cutting plane method for pseudomonotone variational inequalities. *Operations Research Letters*, 20(1):1–6, 1997.

Martin Grötschel, László Lovász, and Alexander Schrijver. *Geometric Algorithms and Combinatorial Optimization*. Springer, 1993.

Guoyong Gu and Junfeng Yang. Tight sublinear convergence rate of the proximal point algorithm for maximal monotone inclusion problems. *SIAM Journal on Optimization*, 30(3):1905–1921, 2020.

Haotian Jiang, Yin Tat Lee, Zhao Song, and Sam Chiu-wai Wong. An improved cutting plane method for convex optimization, convex-concave games, and its applications. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, pages 944–953, 2020.

Donghwan Kim. Accelerated proximal point method for maximally monotone operators. *Mathematical Programming*, 190(1):57–87, 2021.

D Leventhal. Metric subregularity and the proximal point method. *Journal of Mathematical Analysis and Applications*, 360(2):681–688, 2009.

Pierre-Louis Lions and Bertrand Mercier. Splitting algorithms for the sum of two nonlinear operators. *SIAM Journal on Numerical Analysis*, 16(6):964–979, 1979.

Fernando Javier Luque. Asymptotic convergence analysis of the proximal point algorithm. *SIAM Journal on Control and Optimization*, 22(2):277–293, 1984.

Hans-Jakob Lüthi. On the solution of variational inequalities by the ellipsoid method. *Mathematics of Operations Research*, 10(3):515–522, 1985.

George J Minty. Monotone (nonlinear) operators in hilbert space. *Duke Mathematical Journal*, 29(3): 341, 1962.

Jean-Jacques Moreau. Proximité et dualité dans un espace hilbertien. *Bulletin de la Société mathématique de France*, 93:273–299, 1965.

Yu Nesterov and J Ph Vial. Homogeneous analytic center cutting plane methods for convex problems and variational inequalities. *SIAM Journal on Optimization*, 9(3):707–728, 1999.

Yurii Nesterov and Arkadii Nemirovskii. *Interior-point polynomial algorithms in convex programming*. SIAM, 1994.

Jisun Park and Ernest K Ryu. Exact optimal accelerated complexity for fixed-point iterations. In *International Conference on Machine Learning*, pages 17420–17457. PMLR, 2022.

Xiaolong Qin, Shin Min Kang, and Yeol Je Cho. Approximating zeros of monotone operators by proximal point algorithms. *Journal of Global Optimization*, 46(1):75–87, 2010.

R Tyrrell Rockafellar. Monotone operators and the proximal point algorithm. *SIAM journal on control and optimization*, 14(5):877–898, 1976.

Abdurakhmon Sadiev, Laurent Condat, and Peter Richtárik. Stochastic proximal point methods for monotone inclusions under expected similarity. In *OPT 2024: Optimization for Machine Learning*, 2024.

Jie Shen and Li-Ping Pang. A proximal analytic center cutting plane algorithm for solving variational inequality problems. *Journal of Applied Mathematics*, 2012(1):503242, 2012.

Maurice Sion. On general minimax theorems. *Pacific Journal of Mathematics*, 8(1):171–176, 1958.

A Basic Geometry and Preliminary Remarks

We start this section with a fact on basic geometry of the Minty set.

Proposition 20. *The Minty set $\mathcal{K}$ is a compact convex subset of $\mathcal{X}$.*

Proof. For each fixed $y \in \mathcal{X}$ and $v \in A(y)$, the condition $\langle v, y - x\rangle \geq 0$ is equivalent to $\langle v, x\rangle \leq \langle v, y\rangle$, which defines a closed halfspace in $\mathbb{R}^d$. Hence

$$\mathcal{K} = \mathcal{X} \cap \bigcap_{y\in\mathcal{X}} \bigcap_{v\in A(y)} \left\{x \in \mathbb{R}^d : \langle v, x\rangle \leq \langle v, y\rangle\right\}$$

is an intersection of closed convex sets. Since $\mathcal{X}$ is compact, $\mathcal{K}$ is compact as well. □

In the following, we show that arbitrary exact resolvent scales trivialize the approximation problem. Therefore, the realistic regime operates at some constant scale η.

Proposition 21. *Suppose exact constrained resolvent selections are available for every $\eta > 0$. Then one call at the anchor $a = 0$ yields an ε-SVI whenever*

$$\eta \geq \frac{2R^2}{\varepsilon}.$$

Proof. Let (x, u, n) be the exact resolvent output at $a = 0$. Then

$$0 = x + \eta(u + n), \qquad \text{so} \qquad u + n = -\frac{x}{\eta}.$$

Fix $y \in \mathcal{X}$. Since $n \in N_{\mathcal{X}}(x)$,

$$\langle n, y - x\rangle \leq 0 \qquad \Longrightarrow \qquad \langle u, x - y\rangle \leq \langle u + n, x - y\rangle .$$

Using $\|x\| \leq R$ and $\|x - y\| \leq 2R$,

$$\langle u, x - y\rangle \leq \frac{\|x\|}{\eta}\|x - y\| \leq \frac{2R^2}{\eta} \leq \varepsilon.$$

So (x, u) is an ε-SVI. □

Assumption 22 (δ-approximate resolvent oracle)**.** *Fix $\eta > 0$ and $\delta \geq 0$. For every rational anchor $a \in \mathbb{R}^d$, an oracle returns $(\hat{x}, \hat{u}, \hat{n})$ such that*

$$\hat{x} \in \mathcal{X}, \qquad \hat{u} \in A(\hat{x}), \qquad \hat{n} \in N_{\mathcal{X}}(\hat{x}), \qquad \|a - \hat{x} - \eta(\hat{u} + \hat{n})\| \leq \delta.$$

We set

$$\hat{g}(a) := \hat{u} + \hat{n}.$$

B Detailed Proofs

Lemma 23. *Let $s(a) := \|a - J_\eta(a)\|$. Then*

$$\mathrm{Gap}_\eta(a) \leq \frac{1}{\eta}\Big(s(a)^2 + 2Rs(a)\Big).$$

Proof. Let $x^+ = J_\eta(a)$. For any $y \in \mathcal{X}$,

$$\begin{aligned}
\langle g_\eta(a), a - y\rangle &= \frac{1}{\eta}\left\langle a - x^+, a - y\right\rangle \\
&= \frac{1}{\eta}\Big(\left\|a - x^+\right\|^2 + \left\langle a - x^+, x^+ - y\right\rangle\Big) \\
&\leq \frac{1}{\eta}\Big(s(a)^2 + s(a)\left\|x^+ - y\right\|\Big) \\
&\leq \frac{1}{\eta}\Big(s(a)^2 + 2Rs(a)\Big),
\end{aligned}$$

because $x^+, y \in \mathcal{X} \subseteq B_R(0)$. Taking the maximum over $y \in \mathcal{X}$ proves the claim. □

Lemma 8. *If* $\mathrm{Gap}_\eta(a) > \varepsilon$, *then*

$$s(a) > \rho_\eta(\varepsilon, R) := \sqrt{R^2 + \eta\varepsilon} - R.$$

Consequently, the halfspace

$$H(a) := \left\{ z \in \mathbb{R}^d : \langle g_\eta(a), z - a \rangle \le 0 \right\}$$

contains the thickened body $\mathcal{K} + B_{\rho_\eta(\varepsilon,R)}(0)$.

Proof. By Lemma 23,

$$\varepsilon < \mathrm{Gap}_\eta(a) \le \frac{1}{\eta}\Big(s(a)^2 + 2Rs(a)\Big).$$

Hence

$$s(a)^2 + 2Rs(a) - \eta\varepsilon > 0.$$

The positive root of the quadratic polynomial $s^2 + 2Rs - \eta\varepsilon$ is

$$\sqrt{R^2 + \eta\varepsilon} - R.$$

Since $s(a) \ge 0$, we obtain

$$s(a) > \sqrt{R^2 + \eta\varepsilon} - R = \rho_\eta(\varepsilon, R).$$

Now let $z = x + h$ with $x \in \mathcal{K}$ and $\|h\| \le \rho_\eta(\varepsilon, R)$. By Lemma 7,

$$\langle g_\eta(a), x - a \rangle \le -\eta \left\| g_\eta(a) \right\|^2.$$

Therefore

$$\langle g_\eta(a), z - a \rangle = \langle g_\eta(a), x - a \rangle + \langle g_\eta(a), h \rangle \le -\eta \left\| g_\eta(a) \right\|^2 + \rho_\eta(\varepsilon, R) \left\| g_\eta(a) \right\|.$$

Because $s(a) = \eta \left\| g_\eta(a) \right\| > \rho_\eta(\varepsilon, R)$, the right-hand side is strictly negative. Thus z belongs to the halfspace $\{w : \langle g_\eta(a), w - a \rangle \le 0\}$. □

Theorem 9 (Exact resolvent ellipsoid)**.** *Under Assumption 4 and the Minty condition* $\mathcal{K} \neq \varnothing$, *the exact outer method returns an* ε*-SVI after* $\mathcal{O}\Big(d^2 \log \frac{R + \rho_\eta(\varepsilon,R)}{\rho_\eta(\varepsilon,R)}\Big)$ *iterations. Under the nontrivial regime* $\varepsilon \le R^2/\eta$, *this simplifies to*

$$\mathcal{O}\left(d^2 \log \frac{R^2}{\eta\varepsilon}\right).$$

Each iteration uses one exact resolvent call and one linear-optimization call.

Proof. Let

$$C := \mathcal{K} + B_{\rho_\eta(\varepsilon,R)}(0).$$

By Lemma 8, every rejected iterate inserts a central cut containing C. Also

$$C \subseteq B_{R + \rho_\eta(\varepsilon,R)}(0),$$

so the initial ellipsoid contains C. Since $\mathcal{K} \neq \varnothing$, the set C contains a Euclidean ball of radius $\rho_\eta(\varepsilon, R)$. The standard central-cut ellipsoid method [Grötschel et al., 1993] reduces the enclosing volume by a constant factor every iteration. Therefore, after

$$\mathcal{O}\left(d^2 \log \frac{\mathrm{vol}(B_{R+\rho_\eta(\varepsilon,R)}(0))}{\mathrm{vol}(B_{\rho_\eta(\varepsilon,R)}(0))}\right) = \mathcal{O}\left(d^2 \log \frac{R + \rho_\eta(\varepsilon, R)}{\rho_\eta(\varepsilon, R)}\right)$$

iterations, the maintained ellipsoid would have volume smaller than $\mathrm{vol}(C)$ unless the algorithm had terminated. This is impossible because every maintained ellipsoid contains C. Hence the method must terminate earlier.

When it terminates, the gap test has certified $\mathrm{Gap}_\eta(a_t) \le \varepsilon$, so Lemma 6 upgrades the accepted resolvent point to an ε-SVI. Finally,

$$\rho_\eta(\varepsilon, R) = \sqrt{R^2 + \eta\varepsilon} - R = \frac{\eta\varepsilon}{R + \sqrt{R^2 + \eta\varepsilon}}.$$

If in addition $\varepsilon \le R^2/\eta$, then $R + \sqrt{R^2 + \eta\varepsilon} = \Theta(R)$, so

$$\rho_\eta(\varepsilon, R) = \Theta(\eta\varepsilon/R),$$

which yields the simplified $\mathcal{O}(d^2 \log(R^2/(\eta\varepsilon)))$ bound. □

Lemma 24 (Approximate gap certificate)**.** *Assume* (2)*. If*

$$\widehat{\mathrm{Gap}}_\eta(a) \le \frac{\varepsilon}{2} \qquad \textit{and} \qquad \delta \le \frac{\eta\varepsilon}{8R},$$

then $\hat{x}$ is an ε-SVI point.

Proof. Fix $y \in \mathcal{X}$. Since $\hat{n} \in N_{\mathcal{X}}(\hat{x})$,

$$\langle \hat{u}, \hat{x} - y \rangle \le \langle \hat{u} + \hat{n}, \hat{x} - y \rangle = \langle \hat{g}(a), \hat{x} - y \rangle .$$

Set

$$r := a - \hat{x} - \eta\hat{g}(a), \qquad \|r\| \le \delta.$$

Then

$$\begin{aligned}
\langle \hat{g}(a), \hat{x} - y \rangle &= \langle \hat{g}(a), a - y \rangle - \langle \hat{g}(a), a - \hat{x} \rangle \\
&= \langle \hat{g}(a), a - y \rangle - \eta \|\hat{g}(a)\|^2 - \langle \hat{g}(a), r \rangle \\
&\le \widehat{\mathrm{Gap}}_\eta(a) - \eta \|\hat{g}(a)\|^2 + \|\hat{g}(a)\| \, \delta.
\end{aligned}$$

For every $t \ge 0$,

$$-\eta t^2 + \delta t \le \frac{\delta^2}{4\eta}.$$

Hence

$$\langle \hat{u}, \hat{x} - y \rangle \le \frac{\varepsilon}{2} + \frac{\delta^2}{4\eta}.$$

Using $\delta \le \eta\varepsilon/(8R)$ and (2),

$$\frac{\delta^2}{4\eta} \le \frac{\eta\varepsilon^2}{256R^2} \le \frac{\varepsilon}{256} < \frac{\varepsilon}{2}.$$

Therefore $\langle \hat{u}, \hat{x} - y \rangle \le \varepsilon$ for all $y \in \mathcal{X}$, so $\hat{x}$ is an ε-SVI. □

Lemma 25 (Approximate strict separation)**.** *Assume* (2) *and $\delta \le \eta\varepsilon/(16R)$. If*

$$\widehat{\mathrm{Gap}}_\eta(a) > \frac{\varepsilon}{2},$$

then the normalized direction

$$\hat{c}(a) := \frac{\hat{g}(a)}{\|\hat{g}(a)\|}$$

satisfies

$$\langle \hat{c}(a), x - a \rangle \le -\frac{\eta\varepsilon}{16R} \qquad \textit{for all } x \in \mathcal{K}.$$

Consequently, the halfspace

$$\widehat{H}(a) := \left\{ z \in \mathbb{R}^d : \langle \hat{g}(a), z - a \rangle \le 0 \right\}$$

contains

$$\mathcal{K} + B_{\eta\varepsilon/(16R)}(0).$$

Proof. For any $y \in \mathcal{X}$,

$$\langle \hat{g}(a), a - y \rangle = \langle \hat{g}(a), a - \hat{x} \rangle + \langle \hat{g}(a), \hat{x} - y \rangle \le \eta \|\hat{g}(a)\|^2 + \|\hat{g}(a)\| \delta + 2R \|\hat{g}(a)\|.$$

So

$$\widehat{\mathrm{Gap}}_\eta(a) \le \eta \|\hat{g}(a)\|^2 + (2R + \delta) \|\hat{g}(a)\|.$$

If $\eta \|\hat{g}(a)\| \le \eta\varepsilon/(8R)$, then using $\delta \le R$ and (2),

$$\widehat{\mathrm{Gap}}_\eta(a) \le \frac{\eta\varepsilon^2}{64R^2} + \frac{3\varepsilon}{8} \le \frac{\varepsilon}{64} + \frac{3\varepsilon}{8} < \frac{\varepsilon}{2},$$

a contradiction. Hence

$$\eta \|\hat{g}(a)\| > \frac{\eta\varepsilon}{8R}.$$

Now fix $x \in \mathcal{K}$. Because $\hat{u} \in A(\hat{x})$ and $\hat{n} \in N_{\mathcal{X}}(\hat{x})$, the same argument as in Lemma 7 gives

$$\langle \hat{g}(a), \hat{x} - x \rangle \ge 0.$$

Using $a - \hat{x} = \eta \hat{g}(a) + r$,

$$\langle \hat{g}(a), a - x \rangle = \langle \hat{g}(a), a - \hat{x} \rangle + \langle \hat{g}(a), \hat{x} - x \rangle \ge \eta \|\hat{g}(a)\|^2 - \|\hat{g}(a)\| \delta.$$

Divide by $\|\hat{g}(a)\|$ to obtain

$$\langle \hat{c}(a), a - x \rangle \ge \eta \|\hat{g}(a)\| - \delta > \frac{\eta\varepsilon}{8R} - \delta \ge \frac{\eta\varepsilon}{16R}.$$

Equivalently,

$$\langle \hat{c}(a), x - a \rangle \le -\frac{\eta\varepsilon}{16R}.$$

The thickening claim follows by adding a perturbation h with $\|h\| \le \eta\varepsilon/(16R)$ and using Cauchy–Schwarz exactly as in the exact proof. □

Theorem 10 (Robust outer theorem). *Assume the Minty condition, the nontrivial regime*

$$\varepsilon \le \frac{R^2}{\eta},$$

and access to a δ-approximate constrained resolvent oracle at scale η with $\delta \le \frac{\eta\varepsilon}{16R}$. Run the outer method from the initial ball $B_{2R}(0)$, accept whenever $\widehat{\mathrm{Gap}}_\eta(a_t) \le \varepsilon/2$, and otherwise cut with the halfspace $\{z : \langle \hat{g}(a_t), z - a_t \rangle \le 0\}$. Then the method returns an ε-SVI after $\mathcal{O}\left(d^2 \log \frac{R^2}{\eta\varepsilon}\right)$ iterations.

Proof. By Lemma 24, every accepted point is an ε-SVI. Assume an iteration is rejected. Lemma 25 shows that the generated central cut contains

$$\mathcal{K} + B_{\eta\varepsilon/(16R)}(0).$$

This thickened set is contained in the initial ball $B_{2R}(0)$ because (2) implies $\eta\varepsilon/(16R) \le R/16$. Hence every maintained ellipsoid contains a fixed full-dimensional convex body with inscribed radius $\eta\varepsilon/(16R)$. The standard ellipsoid volume argument therefore implies termination after

$$\mathcal{O}\left(d^2 \log \frac{2R}{\eta\varepsilon/(16R)}\right) = \mathcal{O}\left(d^2 \log \frac{R^2}{\eta\varepsilon}\right)$$

iterations. □

Theorem 18 (Converse for weakly monotone). *Fix $\rho \ge 0$ and $\eta > 0$ with $\eta^{-1} > \rho$. Assume A is ρ-hypomonotone and that the constrained resolvent set $J_\eta(a)$ is nonempty. Then $J_\eta(a) = \{x^+\}$ is a singleton. Moreover, if $x \in \mathcal{X}$ is an $\varepsilon_{\mathrm{in}}$-SVI point for the regularized operator $A_{a,\eta}(x) := A(x) + \frac{1}{\eta}(x - a)$, then*

$$\|x - J_\eta(a)\| \le \sqrt{\frac{\varepsilon_{\mathrm{in}}}{\eta^{-1} - \rho}}.$$

Consequently, one recovers a δ-accurate constrained resolvent point by solving the regularized weakly monotone instance to accuracy $\varepsilon_{\mathrm{in}} = (\eta^{-1} - \rho)\delta^2$.

Proof. We first prove uniqueness of the constrained resolvent point. Take any two points $x_1, x_2 \in J_\eta(a)$. Then there exist $u_i \in A(x_i)$ and $n_i \in N_{\mathcal{X}}(x_i)$ such that

$$u_i + n_i + \frac{1}{\eta}(x_i - a) = 0, \qquad i \in \{1, 2\}.$$

Subtract the two relations and take the inner product with $x_1 - x_2$:

$$0 = \langle u_1 - u_2, x_1 - x_2 \rangle + \langle n_1 - n_2, x_1 - x_2 \rangle + \frac{1}{\eta} \|x_1 - x_2\|^2 .$$

Because A is ρ-hypomonotone,

$$\langle u_1 - u_2, x_1 - x_2 \rangle \geq -\rho \|x_1 - x_2\|^2 ,$$

and because the normal cone is monotone,

$$\langle n_1 - n_2, x_1 - x_2 \rangle \geq 0.$$

Therefore

$$0 \geq (\eta^{-1} - \rho) \|x_1 - x_2\|^2 .$$

Since $\eta^{-1} > \rho$, this forces $x_1 = x_2$.

Now let $x \in \mathcal{X}$ be an ε_{in}-SVI point for $A_{a,\eta}$. Then there exists $u \in A(x)$ such that

$$\left\langle u + \frac{1}{\eta}(x - a), x - y \right\rangle \leq \varepsilon_{\text{in}} \qquad \text{for all } y \in \mathcal{X}.$$

Apply this with $y = x^+ = J_\eta(a)$:

$$\left\langle u + \frac{1}{\eta}(x - a), x - x^+ \right\rangle \leq \varepsilon_{\text{in}}.$$

Because $x^+ = J_\eta(a)$, there exist $u^+ \in A(x^+)$ and $n^+ \in N_{\mathcal{X}}(x^+)$ with

$$u^+ + n^+ + \frac{1}{\eta}(x^+ - a) = 0.$$

Since $n^+ \in N_{\mathcal{X}}(x^+)$ and $x \in \mathcal{X}$,

$$\langle n^+, x - x^+ \rangle \leq 0,$$

hence

$$\left\langle u^+ + \frac{1}{\eta}(x^+ - a), x - x^+ \right\rangle \geq 0.$$

Subtracting,

$$\varepsilon_{\text{in}} \geq \langle u - u^+, x - x^+ \rangle + \frac{1}{\eta} \left\|x - x^+\right\|^2 .$$

Hypomonotonicity gives

$$\langle u - u^+, x - x^+ \rangle \geq -\rho \left\|x - x^+\right\|^2 .$$

So

$$\varepsilon_{\text{in}} \geq (\eta^{-1} - \rho) \left\|x - x^+\right\|^2 ,$$

which rearranges to the desired bound. □

C Weak-Minty Proofs and Inexact Variant

We record the exact proof package behind Section 6, together with the robust inexact analogue. We start with a basic result on the geometry of weak-Minty sets.

Proposition 26. *For every $\kappa \geq 0$, the set $\mathcal{K}_\kappa$ is a compact convex subset of $\mathcal{X}$. Moreover $\mathcal{K}_0 = \mathcal{K}$.*

Proof. Fix $\kappa \geq 0$. For each graph point $(y, v) \in \operatorname{gra} A$, the condition

$$\langle v, y - x \rangle \geq -\kappa \operatorname{gap}_A(y, v)^2$$

is equivalent to

$$\langle v, x \rangle \leq \langle v, y \rangle + \kappa \operatorname{gap}_A(y, v)^2,$$

which defines a closed halfspace in $\mathbb{R}^d$. Hence

$$\mathcal{K}_\kappa = \mathcal{X} \cap \bigcap_{(y,v) \in \operatorname{gra} A} \left\{ x \in \mathbb{R}^d : \langle v, x \rangle \leq \langle v, y \rangle + \kappa \operatorname{gap}_A(y, v)^2 \right\}$$

is an intersection of closed convex sets, and is therefore closed and convex. Since $\mathcal{X}$ is compact, $\mathcal{K}_\kappa$ is compact as well. The identity $\mathcal{K}_0 = \mathcal{K}$ is immediate from Definition 3 and Definition 12. □

Lemma 27. *Let $a \in \mathbb{R}^d$, and let (x^+, u^+, n^+) be an exact resolvent triple at a with*

$$g_\eta(a) = u^+ + n^+ = \frac{a - x^+}{\eta}.$$

Then

$$\operatorname{gap}_A(x^+, u^+) \leq 2R \left\| g_\eta(a) \right\|.$$

Proof. Fix any $z \in \mathcal{X}$. Since $n^+ \in N_{\mathcal{X}}(x^+)$ and $z \in \mathcal{X}$, we have

$$\left\langle n^+, z - x^+ \right\rangle \leq 0, \qquad \text{hence} \qquad \left\langle u^+, x^+ - z \right\rangle \leq \left\langle u^+ + n^+, x^+ - z \right\rangle = \left\langle g_\eta(a), x^+ - z \right\rangle.$$

Now $x^+, z \in \mathcal{X} \subseteq B_R(0)$, so $\|x^+ - z\| \leq 2R$. Therefore

$$\left\langle u^+, x^+ - z \right\rangle \leq \left\| g_\eta(a) \right\| \left\| x^+ - z \right\| \leq 2R \left\| g_\eta(a) \right\|.$$

Taking the maximum over $z \in \mathcal{X}$ proves the claim. □

Lemma 28 (Weak-Minty strict separation)**.** *Assume $x \in \mathcal{K}_\kappa$. Then, for every anchor $a \in \mathbb{R}^d$,*

$$\langle g_\eta(a), a - x \rangle \geq (\eta - 4\kappa R^2) \left\| g_\eta(a) \right\|^2.$$

Equivalently, if $g_\eta(a) \neq 0$ and $c(a) := g_\eta(a) / \left\| g_\eta(a) \right\|$, then

$$\langle c(a), x - a \rangle \leq -\lambda_\kappa s(a), \qquad \lambda_\kappa := 1 - \frac{4\kappa R^2}{\eta}, \qquad s(a) := \eta \left\| g_\eta(a) \right\|.$$

Proof. Write $g_\eta(a) = u^+ + n^+$, where $u^+ \in A(x^+)$ and $n^+ \in N_{\mathcal{X}}(x^+)$ for the exact resolvent point $x^+ = J_\eta(a)$. Since $x \in \mathcal{K}_\kappa$ and $(x^+, u^+) \in \operatorname{gra} A$, Definition 12 gives

$$\left\langle u^+, x^+ - x \right\rangle \geq -\kappa \operatorname{gap}_A(x^+, u^+)^2.$$

Since $n^+ \in N_{\mathcal{X}}(x^+)$ and $x \in \mathcal{X}$, we also have

$$\left\langle n^+, x^+ - x \right\rangle \geq 0.$$

Therefore,

$$\left\langle g_\eta(a), x^+ - x \right\rangle = \left\langle u^+, x^+ - x \right\rangle + \left\langle n^+, x^+ - x \right\rangle \geq -\kappa \operatorname{gap}_A(x^+, u^+)^2.$$

Using $a = x^+ + \eta g_\eta(a)$, we obtain

$$\langle g_\eta(a), a - x \rangle = \left\langle g_\eta(a), a - x^+ \right\rangle + \left\langle g_\eta(a), x^+ - x \right\rangle \geq \eta \left\| g_\eta(a) \right\|^2 - \kappa \operatorname{gap}_A(x^+, u^+)^2.$$

Finally, Lemma 27 gives

$$\operatorname{gap}_A(x^+, u^+)^2 \leq 4R^2 \left\| g_\eta(a) \right\|^2,$$

whence

$$\langle g_\eta(a), a - x \rangle \geq (\eta - 4\kappa R^2) \left\| g_\eta(a) \right\|^2.$$

Dividing by $\|g_\eta(a)\|$ yields the normalized form. □

Theorem 13 (Weak-Minty outer theorem)**.** *Let* $\lambda_\kappa := 1 - \frac{4\kappa R^2}{\eta}$. *Assume Assumption 4 and the* κ*-graph weak-Minty condition with* $\lambda_\kappa > 0$. *Then every rejected anchor* a *produces a central cut containing* $\mathcal{K}_\kappa + B_{\lambda_\kappa \rho_\eta(\varepsilon,R)}(0)$. *Consequently, the exact outer method returns an* ε*-SVI after* $\mathcal{O}\left(d^2 \log \frac{R+\lambda_\kappa \rho_\eta(\varepsilon,R)}{\lambda_\kappa \rho_\eta(\varepsilon,R)}\right)$ *iterations. Under the nontrivial regime* $\varepsilon \le R^2/\eta$, *this is* $\mathcal{O}\left(d^2 \log \frac{R^2}{(\eta - 4\kappa R^2)\varepsilon}\right)$.

Proof. By Proposition 26, the set $\mathcal{K}_\kappa$ is a compact convex subset of $\mathcal{X}$. Consider one rejected anchor a in the exact outer method. By Lemma 8, rejection means

$$s(a) = \eta \|g_\eta(a)\| > \rho_\eta(\varepsilon, R).$$

By Lemma 28, for every $x \in \mathcal{K}_\kappa$,

$$\langle c(a), x - a\rangle \le -\lambda_\kappa s(a), \qquad c(a) := \frac{g_\eta(a)}{\|g_\eta(a)\|}.$$

Now fix any $z = x + h$ with $x \in \mathcal{K}_\kappa$ and $\|h\| \le \lambda_\kappa \rho_\eta(\varepsilon, R)$. Then

$$\langle c(a), z - a\rangle = \langle c(a), x - a\rangle + \langle c(a), h\rangle \le -\lambda_\kappa s(a) + \|h\| < -\lambda_\kappa \rho_\eta(\varepsilon, R) + \|h\| \le 0.$$

Hence every rejected cut contains the thickened body

$$\mathcal{K}_\kappa + B_{\lambda_\kappa \rho_\eta(\varepsilon,R)}(0).$$

This body is full-dimensional, convex, and contained in $B_{R+\lambda_\kappa \rho_\eta(\varepsilon,R)}(0)$. Therefore the standard central-cut ellipsoid volume argument [Grötschel et al., 1993] implies termination after

$$\mathcal{O}\left(d^2 \log \frac{R + \lambda_\kappa \rho_\eta(\varepsilon, R)}{\lambda_\kappa \rho_\eta(\varepsilon, R)}\right)$$

rejections. If the method accepts, Lemma 6 returns an ε-SVI. This proves the theorem. □

In the sequel, we prove similar results for the inexact oracle model of Assumption 22.

Lemma 29 (Weak-Minty separation in the inexact oracle model)**.** *Assume* $x \in \mathcal{K}_\kappa$ *and let* $(\hat{x}, \hat{u}, \hat{n})$ *be a* δ*-approximate constrained resolvent triple at anchor* a, *with*

$$\hat{g} := \hat{u} + \hat{n}, \qquad r := a - \hat{x} - \eta \hat{g}, \qquad \|r\| \le \delta.$$

Then

$$\langle \hat{g}, a - x\rangle \ge (\eta - 4\kappa R^2) \|\hat{g}\|^2 - \delta \|\hat{g}\|.$$

Equivalently, if $\hat{g} \ne 0$ *and* $\hat{c} := \hat{g}/\|\hat{g}\|$, *then*

$$\langle \hat{c}, x - a\rangle \le -\lambda_\kappa \hat{s} + \delta, \qquad \hat{s} := \eta \|\hat{g}\|.$$

Proof. Since $x \in \mathcal{K}_\kappa$ and $(\hat{x}, \hat{u}) \in \operatorname{gra} A$, Definition 12 gives

$$\langle \hat{u}, \hat{x} - x\rangle \ge -\kappa \operatorname{gap}_A(\hat{x}, \hat{u})^2.$$

As before, $\hat{n} \in N_{\mathcal{X}}(\hat{x})$ and $x \in \mathcal{X}$ imply

$$\langle \hat{n}, \hat{x} - x\rangle \ge 0,$$

so

$$\langle \hat{g}, \hat{x} - x\rangle \ge -\kappa \operatorname{gap}_A(\hat{x}, \hat{u})^2.$$

By the same proof as Lemma 27,

$$\operatorname{gap}_A(\hat{x}, \hat{u}) \le 2R \|\hat{g}\|.$$

Now use $a = \hat{x} + \eta \hat{g} + r$:

$$\begin{aligned}\langle \hat{g}, a - x\rangle &= \eta \|\hat{g}\|^2 + \langle \hat{g}, r\rangle + \langle \hat{g}, \hat{x} - x\rangle \\ &\ge \eta \|\hat{g}\|^2 - \delta \|\hat{g}\| - \kappa \operatorname{gap}_A(\hat{x}, \hat{u})^2 \\ &\ge (\eta - 4\kappa R^2) \|\hat{g}\|^2 - \delta \|\hat{g}\|.\end{aligned}$$

Dividing by $\|\hat{g}\|$ gives the normalized statement. □

Proposition 30. *Assume the κ-graph weak-Minty condition with $\lambda_\kappa > 0$, the nontrivial regime $\varepsilon \le R^2/\eta$, and access to a δ-approximate constrained resolvent oracle with*

$$\delta \le \frac{\lambda_\kappa \eta \varepsilon}{16R}.$$

Run the inexact outer method from Theorem 10, accepting whenever $\widehat{\mathrm{Gap}}_\eta(a_t) \le \varepsilon/2$ and otherwise cutting with $\{z : \langle \hat{g}(a_t), z - a_t\rangle \le 0\}$. Then the method returns an ε-SVI after

$$\mathcal{O}\left(d^2 \log \frac{R^2}{\lambda_\kappa \eta \varepsilon}\right)$$

iterations.

Proof. If the method accepts, Theorem 10 already yields an ε-SVI because

$$\delta \le \frac{\lambda_\kappa \eta \varepsilon}{16R} \le \frac{\eta\varepsilon}{16R} < \frac{\eta\varepsilon}{8R}.$$

Now consider a rejected anchor a. Write $\hat{g} := \hat{g}(a)$ and $\hat{s} := \eta \|\hat{g}\|$. We first claim that rejection implies

$$\hat{s} > \frac{\eta\varepsilon}{8R}.$$

Indeed, if $\hat{s} \le \eta\varepsilon/(8R)$, then

$$\widehat{\mathrm{Gap}}_\eta(a) = \max_{y\in\mathcal{X}} \langle \hat{g}, a - y\rangle \le \eta \|\hat{g}\|^2 + \delta \|\hat{g}\| + 2R\|\hat{g}\| = \frac{1}{\eta}\big(\hat{s}^2 + (2R+\delta)\hat{s}\big).$$

Using $\delta \le \lambda_\kappa\eta\varepsilon/(16R) \le \eta\varepsilon/(16R) \le R/16$ and $\varepsilon \le R^2/\eta$, we obtain

$$\widehat{\mathrm{Gap}}_\eta(a) \le \frac{\eta\varepsilon^2}{64R^2} + \frac{(2R + R/16)\varepsilon}{8R} \le \frac{\varepsilon}{64} + \frac{33\varepsilon}{128} < \frac{\varepsilon}{2},$$

contradicting rejection. Hence $\hat{s} > \eta\varepsilon/(8R)$, as claimed.

Now fix any $x \in \mathcal{K}_\kappa$. By Lemma 29,

$$\langle \hat{g}, a - x\rangle \ge \lambda_\kappa \hat{s}\, \|\hat{g}\| - \delta \|\hat{g}\|,$$

and therefore, for $\hat{c} := \hat{g}/\|\hat{g}\|$,

$$\langle \hat{c}, a - x\rangle \ge \lambda_\kappa \hat{s} - \delta > \lambda_\kappa \frac{\eta\varepsilon}{8R} - \delta \ge \frac{\lambda_\kappa\eta\varepsilon}{16R}.$$

Equivalently,

$$\langle \hat{c}, x - a\rangle \le -\frac{\lambda_\kappa\eta\varepsilon}{16R}.$$

Thus every rejected cut contains the thickened body

$$\mathcal{K}_\kappa + B_{\lambda_\kappa\eta\varepsilon/(16R)}(0).$$

Applying the standard central-cut ellipsoid volume argument completes the proof. □

D Certificate Proofs

The following theorem shows that, under the inexact oracle model, the outer method either finds an approximate SVI solution or returns a strict GEVI certificate.

Theorem 31. *Assume the hypotheses of Theorem 10 and set*

$$\mu := \frac{\eta\varepsilon}{16R}, \qquad \widehat{\gamma}_{\mathrm{cert}} := \frac{\mu^2}{2\eta} = \frac{\eta\varepsilon^2}{512R^2}.$$

Then the inexact outer method returns, after

$$\mathcal{O}\left(d^2 \log \frac{R^2}{\eta\varepsilon}\right)$$

iterations and polynomial additional time, either

(i) an ε-SVI point, or

(ii) a $\widehat{\gamma}_{\mathrm{cert}}$-strict GEVI certificate supported on at most $d+2$ graph points.

Proof. If the method accepts, Theorem 10 already guarantees an ε-SVI. Suppose instead that every iterate is rejected, and let $(a_t, \hat{x}_t, \hat{u}_t, \hat{n}_t, \hat{g}_t, r_t)_{t=1}^{T}$ be the inexact transcript, where

$$r_t := a_t - \hat{x}_t - \eta \hat{g}_t, \qquad \|r_t\| \le \delta.$$

Fix $x \in \mathcal{X}$ and suppose that

$$\langle \hat{u}_t, x - \hat{x}_t \rangle < \widehat{\gamma}_{\mathrm{cert}} \qquad \text{for every } t \in [T].$$

Define $\hat{c}_t := \hat{g}_t / \|\hat{g}_t\|$. By Lemma 25,

$$\eta \|\hat{g}_t\| > \frac{\eta\varepsilon}{8R} = 2\mu.$$

Using $\hat{n}_t \in N_{\mathcal{X}}(\hat{x}_t)$ and $x \in \mathcal{X}$,

$$\begin{aligned}
\langle \hat{c}_t, x - a_t \rangle &= \frac{1}{\|\hat{g}_t\|}\Big(\langle \hat{g}_t, x - \hat{x}_t \rangle - \eta \|\hat{g}_t\|^2 - \langle \hat{g}_t, r_t \rangle \Big) \\
&= \frac{1}{\|\hat{g}_t\|}\Big(\langle \hat{u}_t, x - \hat{x}_t \rangle + \langle \hat{n}_t, x - \hat{x}_t \rangle - \eta \|\hat{g}_t\|^2 - \langle \hat{g}_t, r_t \rangle \Big) \\
&\le \frac{\widehat{\gamma}_{\mathrm{cert}}}{\|\hat{g}_t\|} - \eta \|\hat{g}_t\| + \delta.
\end{aligned}$$

Now

$$\frac{\widehat{\gamma}_{\mathrm{cert}}}{\|\hat{g}_t\|} < \frac{\mu^2}{2\eta} \cdot \frac{\eta}{2\mu} = \frac{\mu}{4},$$

and $\delta \le \mu$ by Theorem 10. Hence

$$\langle \hat{c}_t, x - a_t \rangle < -\frac{3\mu}{4}.$$

Therefore the whole ball $x + B_{3\mu/4}(0)$ lies in every rejected central cut. If the algorithm were run until the final ellipsoid has volume smaller than $\mathrm{vol}(B_{3\mu/4}(0))$, this would be impossible. Thus for every $x \in \mathcal{X}$,

$$\max_{1 \le t \le T} \langle \hat{u}_t, x - \hat{x}_t \rangle \ge \widehat{\gamma}_{\mathrm{cert}}.$$

Equivalently,

$$\min_{x \in \mathcal{X}} \max_{1 \le t \le T} \langle \hat{u}_t, x - \hat{x}_t \rangle \ge \widehat{\gamma}_{\mathrm{cert}}.$$

Now define

$$L(\lambda, x) := \sum_{t=1}^{T} \lambda_t \langle \hat{u}_t, x - \hat{x}_t \rangle, \qquad \lambda \in \Delta_T.$$

For each fixed x,

$$\max_{\lambda \in \Delta_T} L(\lambda, x) = \max_{1 \le t \le T} \langle \hat{u}_t, x - \hat{x}_t \rangle.$$

Hence

$$\min_{x \in \mathcal{X}} \max_{\lambda \in \Delta_T} L(\lambda, x) \ge \widehat{\gamma}_{\mathrm{cert}}.$$

By Sion's minimax theorem [Sion, 1958],

$$\max_{\lambda \in \Delta_T} \min_{x \in \mathcal{X}} L(\lambda, x) \ge \widehat{\gamma}_{\mathrm{cert}}.$$

Thus there exists $\lambda \in \Delta_T$ such that

$$\sum_{t=1}^{T} \lambda_t \langle \hat{u}_t, x - \hat{x}_t \rangle \ge \widehat{\gamma}_{\mathrm{cert}} \qquad \text{for all } x \in \mathcal{X},$$

which is exactly a strict GEVI certificate. Support compression to at most $d+2$ graph points follows from Appendix F. □

E Cutting-plane Upgrade

The strict Yosida cut is a genuine solution-or-separation oracle for a thickened Minty body, so one can replace the outer ellipsoid by the more recent cutting-plane method of Jiang et al. [2020].

Theorem 32. *Under the hypotheses of Theorem 9, one can replace the outer ellipsoid by the cutting-plane method of Jiang et al. [2020]. The resulting algorithm returns an ε-SVI using*

$$\mathcal{O}\left(d \log \frac{d(R+\rho_\eta(\varepsilon,R))}{\rho_\eta(\varepsilon,R)}\right)$$

exact resolvent calls and linear-optimization calls. Under the nontrivial regime $\varepsilon \le R^2/\eta$, this simplifies to

$$\mathcal{O}\left(d \log \frac{dR^2}{\eta\varepsilon}\right).$$

Proof. Lemma 8 shows that the exact Yosida-gap routine is a solution-or-separation oracle for the convex body

$$\mathcal{K} + B_{\rho_\eta(\varepsilon,R)}(0) \subseteq B_{R+\rho_\eta(\varepsilon,R)}(0),$$

which contains a Euclidean ball of radius $\rho_\eta(\varepsilon,R)$. Applying Jiang et al. [2020] yields the stated oracle complexity. □

Theorem 33. *Under the hypotheses of Theorem 10, one can replace the outer ellipsoid by the cutting-plane method of Jiang et al. [2020]. The resulting algorithm returns an ε-SVI using*

$$\mathcal{O}\left(d \log \frac{dR^2}{\eta\varepsilon}\right)$$

approximate resolvent calls and linear-optimization calls.

Proof. Lemma 25 shows that the inexact rejected-gap routine is a solution-or-separation oracle for

$$\mathcal{K} + B_{\eta\varepsilon/(32R)}(0),$$

which is contained in $B_{2R}(0)$ under (2). Applying Jiang et al. [2020] gives the claim. □

F Exact Duality and Support Compression

First, we show that any GEVI certificate can be compressed to a small support of at most $d+2$ graph points by Carathéodory's theorem.

Proposition 34. *Any γ-strict GEVI certificate can be replaced, with the same strictness γ, by another γ-strict GEVI certificate supported on at most $d+2$ graph points.*

Proof. Consider a certificate

$$\mu = \sum_{t=1}^{T} \lambda_t \, \delta_{(x_t,u_t)}.$$

Its defining inequality is

$$\sum_{t=1}^{T} \lambda_t \, \langle u_t, x - x_t \rangle \ge \gamma \qquad \text{for all } x \in \mathcal{X}.$$

Introduce

$$\bar{u} := \sum_{t=1}^{T} \lambda_t u_t \in \mathbb{R}^d, \qquad \bar{b} := \sum_{t=1}^{T} \lambda_t \, \langle u_t, x_t \rangle \in \mathbb{R}.$$

Then the certificate inequality is equivalent to

$$\langle \bar{u}, x \rangle - \bar{b} \geq \gamma \qquad \text{for all } x \in \mathcal{X}.$$

So the certificate depends only on the point

$$(\bar{u}, \bar{b}) = \sum_{t=1}^{T} \lambda_t (u_t, \langle u_t, x_t \rangle) \in \mathbb{R}^{d+1}.$$

By Carathéodory's theorem in $\mathbb{R}^{d+1}$, the same point can be represented as a convex combination of at most $d+2$ of the vectors $(u_t, \langle u_t, x_t \rangle)$. Using those new weights on the corresponding graph points yields a certificate with the same strictness. □

Assumption 35. *The graph*

$$\Gamma := \operatorname{gra} A = \{(y, u) \in \mathcal{X} \times \mathbb{R}^d : u \in A(y)\}$$

is compact.

We next establish the exact duality between Minty feasibility and strict GEVI certificates, showing that infeasibility is witnessed by a finite-support certificate.

Theorem 36. *Assume Assumption 35 and define*

$$\gamma^\star(A, \mathcal{X}) := \min_{x \in \mathcal{X}} \max_{(y,u) \in \Gamma} \langle u, x - y \rangle .$$

Then

$$\gamma^\star(A, \mathcal{X}) = \max_{\mu \in \mathcal{P}(\Gamma)} \min_{x \in \mathcal{X}} \int_\Gamma \langle u, x - y \rangle \; d\mu(y, u).$$

Moreover, the following are equivalent:

(i) $\mathcal{K} \neq \varnothing$*;*

(ii) $\gamma^\star(A, \mathcal{X}) = 0$*;*

(iii) there is no γ*-strict GEVI certificate with* $\gamma > 0$*.*

When $\mathcal{K} = \varnothing$*, there exists a* $\gamma^\star(A, \mathcal{X})$*-strict GEVI certificate, and one may choose it with support at most* $d + 2$*.*

Proof. Because Γ and $\mathcal{X}$ are compact and the payoff $(\mu, x) \mapsto \int_\Gamma \langle u, x - y \rangle \; d\mu(y, u)$ is bilinear and continuous, Sion's theorem gives the minimax identity. Now

$$\mathcal{K} \neq \varnothing \iff \exists x \in \mathcal{X} \text{ s.t. } \langle u, x - y \rangle \leq 0 \text{ for all } (y, u) \in \Gamma \iff \gamma^\star(A, \mathcal{X}) \leq 0.$$

Since taking $x = y$ shows the inner maximum is always nonnegative, $\gamma^\star(A, \mathcal{X}) \geq 0$. Hence $\gamma^\star(A, \mathcal{X}) = 0$ exactly when $\mathcal{K} \neq \varnothing$. If $\gamma^\star(A, \mathcal{X}) > 0$, an optimal measure in the minimax problem is a strict GEVI certificate, and Proposition 34 gives the support bound. □

G Oracle Realizations

We present representative settings in which the fixed-scale constrained resolvent oracle assumed in the main text can be implemented efficiently. These examples illustrate how the required oracle access can be realized for structured inclusion problems.

G.1 Canonical Realizations Beyond Point Evaluation

Proposition 37 (Convex feasibility via normal cones)**.** *Let* $C \subseteq \mathbb{R}^d$ *be closed and convex and define the set-valued operator*

$$A(x) := \begin{cases} N_C(x), & x \in C \cap \mathcal{X}, \\ \varnothing, & x \in \mathcal{X} \setminus C. \end{cases}$$

Then every ε-SVI point of A belongs to $C \cap \mathcal{X}$. If moreover the qualification

$$N_{C\cap\mathcal{X}}(x) = N_C(x) + N_{\mathcal{X}}(x) \qquad \text{for all } x \in C \cap \mathcal{X}$$

holds, then for every anchor $a \in \mathbb{R}^d$ and every scale $\eta > 0$,

$$J_\eta(a) = P_{C\cap\mathcal{X}}(a).$$

Consequently, fixed-scale resolvent access reduces to Euclidean projection onto the feasible set.

Proof. If x is an ε-SVI point, then by Definition 1 there exists $u \in A(x)$; by construction of A, this forces $x \in C \cap \mathcal{X}$. Conversely, any $x \in C \cap \mathcal{X}$ is a 0-SVI point because $0 \in N_C(x)$. For the resolvent identity, $x \in J_\eta(a)$ means precisely that there exist $u \in N_C(x)$ and $n \in N_{\mathcal{X}}(x)$ with

$$a = x + \eta(u + n).$$

Under the qualification, this is equivalent to

$$a - x \in \eta N_{C\cap\mathcal{X}}(x),$$

which is the standard optimality condition for the Euclidean projection of a onto $C \cap \mathcal{X}$. □

Proposition 38 (Composite convex optimization)**.** *Take $A = \partial f$ for a proper, closed, convex function f. Then the constrained resolvent is exactly*

$$J_\eta(a) = \operatorname{argmin}_{x\in\mathcal{X}} \left\{ f(x) + \frac{1}{2\eta} \|x - a\|^2 \right\}.$$

Hence, if this strongly convex subproblem can be solved to graph residual δ in time $\mathrm{poly}(d, \log(1/\delta))$, then one can compute $x \in \mathcal{X}$ with

$$f(x) \le \min_{y\in\mathcal{X}} f(y) + \varepsilon$$

in time $\mathrm{poly}(d, \log(1/\varepsilon))$.

Proof. The first-order optimality condition for the displayed problem is

$$0 \in \partial f(x) + N_{\mathcal{X}}(x) + \frac{1}{\eta}(x - a),$$

which is equivalent to

$$a \in x + \eta(\partial f(x) + N_{\mathcal{X}}(x)).$$

So its minimizer is exactly the constrained resolvent point. The outer claim then follows from Corollary 11 and the subgradient inequality. □

Proposition 39 (Compact primal-dual saddle/KKT systems)**.** *Let $X \subseteq \mathbb{R}^p$ and $Y \subseteq \mathbb{R}^q$ be compact convex sets, let $K : \mathbb{R}^p \to \mathbb{R}^q$ be linear, let $f : X \to (-\infty, +\infty]$ and $g : Y \to (-\infty, +\infty]$ be proper, closed, and convex, and consider the monotone operator on $Z := X \times Y$,*

$$A(x, y) := \big(\partial f(x) + K^\top y,\ \partial g(y) - Kx\big).$$

Then the constrained resolvent of $A + N_Z$ at anchor $a = (a_x, a_y)$ is the saddle point of

$$\min_{x\in X} \max_{y\in Y} \left\{ f(x) + \langle Kx, y\rangle - g(y) + \frac{1}{2\eta} \|x - a_x\|^2 - \frac{1}{2\eta} \|y - a_y\|^2 \right\}.$$

Hence, if this strongly regularized saddle problem can be solved to graph residual δ in time $\mathrm{poly}(d, \log(1/\delta))$, then the associated Minty-SVI problem is solvable in $\mathrm{poly}(d, \log(1/\varepsilon))$ time.

Proof. The KKT conditions of the displayed saddle problem are

$$0 \in \partial f(x^+) + K^\top y^+ + N_X(x^+) + \frac{1}{\eta}(x^+ - a_x),$$

and

$$0 \in \partial g(y^+) - Kx^+ + N_Y(y^+) + \frac{1}{\eta}(y^+ - a_y).$$

Together they are equivalent to

$$a \in (x^+, y^+) + \eta(A(x^+, y^+) + N_Z(x^+, y^+)).$$

So $(x^+, y^+) = J_\eta(a)$, and Corollary 11 gives the runtime statement. □

These constructions also extend to symmetric monotone quadratic and complementarity systems, models with closed-form proximal maps such as soft-thresholding and singular-value thresholding, and strongly regularized cone programs.

Theorem 40. *Consider any of the anchored subproblems in Propositions 37 to 39 at a fixed scale $\eta > 0$. Assume that, for every anchor a, the corresponding subproblem admits a polynomial-size LP, SOCP, SDP, or more generally self-concordant conic representation with rational data depending affinely on a. Then standard path-following interior-point methods compute a primal-dual point with KKT residual at most δ in time $\mathrm{poly}(\mathrm{size}(\mathrm{input}), d, \log(1/\delta))$ [Nesterov and Nemirovskii, 1994]. Consequently, the associated constrained resolvent can be computed to graph residual tolerance δ in the same complexity, and Corollary 11 yields a Minty-SVI algorithm with runtime $\mathrm{poly}(\mathrm{size}(\mathrm{input}), d, \log(1/\varepsilon))$.*

Proof. In each case, the resolvent point is characterized by a strongly regularized conic KKT system. A primal-dual iterate whose KKT residual is at most δ furnishes a graph-residual triple $(\hat{x}, \hat{u}, \hat{n})$ with

$$\|a - \hat{x} - \eta(\hat{u} + \hat{n})\| \le \delta.$$

Standard interior-point theory yields such a point in time polynomial in the input size and logarithmic in $1/\delta$. Corollary 11 then gives the claimed outer runtime. □

Remark 41. *In these examples, the constrained resolvent can be computed through projections, proximal evaluations, or regularized KKT solves. These oracle realizations extend beyond the single-valued Lipschitz point-evaluation framework considered by Anagnostides et al. [2025], allowing the localization method to apply directly to structured set-valued inclusions.*

G.2 Bridge Back to Single-Valued Lipschitz VIs

We now connect fixed-scale constrained resolvent access to the classical point-evaluation model for single-valued Lipschitz VIs. The following proposition characterizes the constrained resolvent through an anchored Stampacchia VI. Under hypomonotonicity, a suitable regularization parameter makes the anchored operator strongly monotone, which enables efficient resolvent computation using existing VI methods.

Proposition 42 (Single-valued anchored regularization)**.** *Assume $A(x) = \{F(x)\}$ with single-valued $F : \mathcal{X} \to \mathbb{R}^d$. For anchor a and scale $\eta > 0$, a point $x^+ \in \mathcal{X}$ belongs to $J_\eta(a)$ if and only if*

$$\langle F(x^+) + \eta^{-1}(x^+ - a), y - x^+ \rangle \ge 0 \qquad \text{for all } y \in \mathcal{X}.$$

Equivalently, x^+ is an exact SVI point of the anchored field

$$F_{a,\eta}(x) := F(x) + \eta^{-1}(x - a).$$

If moreover F is ρ-hypomonotone and $\eta^{-1} > \rho$, then $F_{a,\eta}$ is $(\eta^{-1} - \rho)$-strongly monotone. Consequently, any $\varepsilon_{\mathrm{in}}$-SVI point of the anchored instance is within

$$\sqrt{\frac{\varepsilon_{\mathrm{in}}}{\eta^{-1} - \rho}}$$

of the exact resolvent point $J_\eta(a)$.

Proof. If $x^+ \in J_\eta(a)$, then by Definition 2 there exists $n^+ \in N_{\mathcal{X}}(x^+)$ with

$$0 = F(x^+) + n^+ + \eta^{-1}(x^+ - a).$$

Since $n^+ \in N_{\mathcal{X}}(x^+)$, this is equivalent to the displayed variational inequality. The converse is the same argument in reverse. If F is ρ-hypomonotone, then for all $x, y \in \mathcal{X}$,

$$\langle F_{a,\eta}(x) - F_{a,\eta}(y), x - y \rangle = \langle F(x) - F(y), x - y \rangle + \eta^{-1}\|x - y\|^2 \ge (\eta^{-1} - \rho)\|x - y\|^2.$$

So $F_{a,\eta}$ is strongly monotone. The final approximation claim is exactly Theorem 18 specialized to the single-valued case. □

Corollary 43. *Let $F : \mathcal{X} \to \mathbb{R}^d$ be single-valued, L-Lipschitz, and ρ-hypomonotone. If $\eta^{-1} > \rho$, then for every anchor a one can compute $J_\eta(a)$ to Euclidean error δ in time*

$$\mathrm{poly}\big(d, \log(L + \eta^{-1}), \log(1/\delta)\big)$$

by running the ellipsoid of Lüthi [1985] (for strongly monotone VIs) on the anchored field $F_{a,\eta}$ with target accuracy

$$\varepsilon_{\mathrm{in}} := (\eta^{-1} - \rho)\delta^2.$$

Proof. By Proposition 42, computing the resolvent is equivalent to solving the anchored SVI. The field $F_{a,\eta}$ is $(L + \eta^{-1})$-Lipschitz and strongly monotone, hence Minty. Therefore Anagnostides et al. [2025] compute an $\varepsilon_{\mathrm{in}}$-SVI of the anchored instance in polylogarithmic accuracy. Theorem 18 then converts that inner SVI tolerance into Euclidean resolvent error δ. □

As a simple consequence, every single-valued Lipschitz VI admits efficient fixed-scale resolvent computation, since Lipschitzness alone implies hypomonotonicity.

Corollary 44. *Assume $F : \mathcal{X} \to \mathbb{R}^d$ is single-valued and L-Lipschitz. Then F is automatically L-hypomonotone. Consequently, for every anchor a and every scale $\eta \in (0, 1/L)$, one can compute the resolvent point $J_\eta(a)$ to Euclidean error δ in time*

$$\mathrm{poly}\big(d, \log(L + \eta^{-1}), \log(1/\delta)\big).$$

In particular, taking $\eta = 1/(2L)$ gives a fixed-scale resolvent-point routine with polylogarithmic dependence on the target error.

Proof. For any $x, y \in \mathcal{X}$,

$$\langle F(x) - F(y), x - y\rangle \geq -\|F(x) - F(y)\| \, \|x - y\| \geq -L\|x - y\|^2,$$

so F is L-hypomonotone. Now apply Corollary 43 with $\rho = L$. □

From point error to graph residual. The Euclidean approximation guarantee in Corollary 44 controls the distance to an exact resolvent point but does not ensure the graph-residual accuracy required by Corollary 11. This distinction arises because the normal-cone mapping $N_{\mathcal{X}}$ can be discontinuous at the boundary of $\mathcal{X}$. The following corollary shows how to obtain the required graph-residual guarantee when Euclidean projection onto $\mathcal{X}$ is available.

Corollary 45. *Let $F : \mathcal{X} \to \mathbb{R}^d$ be a single-valued L-Lipschitz operator, and fix $\eta \in (0, 1/L)$. Suppose Euclidean projection onto $\mathcal{X}$ can be computed in time $T_{\mathrm{proj}}(d)$. For an anchor a and a point $\tilde{x} \in \mathcal{X}$, define*

$$\bar{x} := P_{\mathcal{X}}\big(a - \eta F(\tilde{x})\big), \qquad \bar{n} := \frac{a - \eta F(\tilde{x}) - \bar{x}}{\eta}, \qquad \bar{u} := F(\bar{x}).$$

Writing $x^+ = J_\eta(a)$, the resulting triple satisfies $\bar{x} \in \mathcal{X}$, $\bar{u} = F(\bar{x})$, $\bar{n} \in N_{\mathcal{X}}(\bar{x})$, and

$$\|a - \bar{x} - \eta(\bar{u} + \bar{n})\| \leq \eta L(1 + \eta L)\left\|\tilde{x} - x^+\right\|.$$

Consequently, combining Corollary 44 with one projection yields a δ-approximate constrained resolvent selection in time

$$\mathrm{poly}\big(d, \log(L + \eta^{-1}), \log(1/\delta)\big) + T_{\mathrm{proj}}(d).$$

For single-valued Lipschitz Minty VIs, this provides the oracle access required by Corollary 11. Moreover, the strict GEVI certificates of Theorem 31 reduce to strict EVI certificates in this setting.

If the instance also satisfies the constrained weak-Minty condition of Anagnostides et al. [2025] with

$$\kappa \leq C\frac{L}{(B + 4RL)^2}, \qquad C < 1,$$

where $B := \sup_{x \in \mathcal{X}} \|F(x)\|$, then Proposition 30 applies with $\eta = 1/(2L)$.

Proof. The projection optimality condition gives $\bar{x} \in \mathcal{X}$ and $\bar{n} \in N_{\mathcal{X}}(\bar{x})$, while $\bar{u} = F(\bar{x})$. By Proposition 42, the exact resolvent point satisfies

$$x^+ = P_{\mathcal{X}}\big(a - \eta F(x^+)\big).$$

The nonexpansiveness of Euclidean projection and the Lipschitz continuity of F imply

$$\left\|\bar{x} - x^+\right\| \le \eta \left\|F(\tilde{x}) - F(x^+)\right\| \le \eta L \left\|\tilde{x} - x^+\right\|.$$

By the definitions of $\bar{u}$ and $\bar{n}$,

$$a - \bar{x} - \eta(\bar{u} + \bar{n}) = \eta\big(F(\tilde{x}) - F(\bar{x})\big).$$

Hence,

$$\begin{aligned}\|a - \bar{x} - \eta(\bar{u} + \bar{n})\| &\le \eta L \|\tilde{x} - \bar{x}\| \\ &\le \eta L(1 + \eta L) \left\|\tilde{x} - x^+\right\|.\end{aligned}$$

Therefore, a point approximation satisfying

$$\left\|\tilde{x} - x^+\right\| \le \frac{\delta}{\eta L(1 + \eta L)}$$

produces a δ-approximate constrained resolvent triple after one projection. The runtime follows from Corollary 44.

The Minty-SVI guarantee follows from Corollary 11. Since $A(x) = \{F(x)\}$, the strict GEVI alternative of Theorem 31 coincides with the strict EVI alternative.

For the weak-Minty claim, using $B \ge 0$ and $\eta = 1/(2L)$ gives

$$\kappa \le \frac{CL}{(B + 4RL)^2} \le \frac{C}{16LR^2} < \frac{1}{8LR^2} = \frac{\eta}{4R^2}.$$

Thus the admissibility condition of Proposition 30 is satisfied. □

The two-player smooth concave-game construction in Anagnostides et al. [2025, Section 5] also leads to a single-valued Lipschitz VI on a compact truncated domain. When Euclidean projection onto this domain is efficient, the same argument provides the required graph-residual resolvent oracle. Furthermore, Anagnostides et al. [2025, Lemma 5.3] show how sufficiently strict EVI certificates for the transformed VI yield strict coarse correlated equilibria.

H Why the Direct Converse Fails Under Minty

The forward reduction established in Corollary 11 holds under the Minty condition alone and can be summarized as

$$\text{efficient approximate resolvent} \implies \text{polylog-accuracy Minty-SVI}.$$

The reverse reduction requires additional structure. In particular, applying a Minty-SVI solver to recover $J_\eta(a)$ requires solving the anchored problem associated with

$$A_{a,\eta}(x) := A(x) + \frac{1}{\eta}(x - a).$$

Under the Minty condition alone, this regularization may have an empty Minty set even when $J_\eta(a)$ is nonempty. The following proposition gives an explicit counterexample.

Proposition 46. *There exist a compact convex set $\mathcal{X} \subset \mathbb{R}$, a single-valued affine operator $A : \mathcal{X} \to \mathbb{R}$, and parameters $\eta > 0$ and $a \in \mathbb{R}$ such that*

1. *A has a nonempty Minty set on $\mathcal{X}$;*
2. *the regularized operator $A_{a,\eta}(x) := A(x) + \eta^{-1}(x - a)$ has an empty Minty set on $\mathcal{X}$;*
3. *the constrained resolvent set $J_\eta(a)$ is nonempty.*

Proof. Consider

$$\mathcal{X} = [-1, 1], \qquad A(x) = -2x - 3, \qquad \eta = 1, \qquad a = -3.$$

Since $A(y) < 0$ for every $y \in \mathcal{X}$, the point $x^\star = 1$ satisfies

$$A(y)(y - 1) \geq 0 \qquad \text{for all } y \in \mathcal{X}.$$

Conversely, for any $x < 1$, choosing $y = 1$ gives

$$A(1)(1 - x) = -5(1 - x) < 0.$$

Hence the Minty set of A is $K = \{1\}$.

For the chosen anchor and scale, the regularized operator becomes

$$A_{a,1}(x) = A(x) + (x - a) = -x.$$

If $x \in \mathcal{X}$ satisfied the Minty condition for $A_{a,1}$, evaluating the Minty inequality at $y = 1$ and $y = -1$ would require

$$(-1)(1 - x) \geq 0 \quad \Longrightarrow \quad x \geq 1,$$

and

$$(1)(-1 - x) \geq 0 \quad \Longrightarrow \quad x \leq -1.$$

These inequalities are incompatible, so the Minty set of $A_{a,1}$ is empty.

Nevertheless, the constrained resolvent is nonempty. Indeed, the relation

$$a = x + A(x) + n, \qquad n \in N_{\mathcal{X}}(x),$$

reduces to

$$-3 = x - 2x - 3 + n, \qquad \text{or equivalently} \qquad n = x.$$

On $\mathcal{X} = [-1, 1]$, this condition holds precisely at $x \in \{-1, 0, 1\}$. Therefore,

$$J_1(-3) = \{-1, 0, 1\} \neq \emptyset.$$

□

The example shows that Minty feasibility of A does not guarantee Minty feasibility of its anchored regularization. Consequently, a Minty-SVI solver cannot be applied directly to the regularized subproblem without additional assumptions.

I Linear Convergence under a Yosida Error Bound

Our ellipsoid method achieves polylogarithmic oracle complexity under the Minty condition alone, without requiring additional regularity. For monotone inclusions satisfying a Yosida error bound, however, the simpler fixed-scale resolvent iteration already converges linearly, making the cutting-plane procedure unnecessary.

Definition 47 (Yosida error bound)**.** *A monotone inclusion satisfies a* Yosida error bound *at scale η with parameter $\kappa \in (0, 1]$ if*

$$\eta \, \|g_\eta(a)\| \geq \kappa \, \operatorname{dist}(a, \mathcal{K}) \qquad \textit{for all } a \in \mathcal{X}.$$

Theorem 48. *Assume A is monotone and satisfies the Yosida error bound at scale η with parameter $\kappa \in (0, 1]$. Starting from any $a_0 \in \mathcal{X}$, consider the exact resolvent iteration*

$$a_{t+1} := J_\eta(a_t).$$

Then

$$\operatorname{dist}(a_{t+1}, \mathcal{K})^2 \leq (1 - \kappa^2) \operatorname{dist}(a_t, \mathcal{K})^2 \qquad \textit{for all } t \geq 0.$$

Consequently, an ε-SVI can be computed in

$$\mathcal{O}\left(1 + \kappa^{-2} \log\left(1 + \frac{R + \rho_\eta(\varepsilon, R)}{\rho_\eta(\varepsilon, R)}\right)\right)$$

exact resolvent evaluations. In the regime $\varepsilon \leq R^2/\eta$, this simplifies to

$$\mathcal{O}\left(\kappa^{-2} \log\left(1 + \frac{R^2}{\eta\varepsilon}\right)\right).$$

For monotone A, the returned point also satisfies the ε-Minty inequalities.

Proof. Let $x_t^\star$ be a nearest Minty point to a_t. By the resolvent identity and Lemma 7,

$$\begin{aligned}\operatorname{dist}(a_{t+1},\mathcal{K})^2 &\le \|a_t - x_t^\star - \eta g_\eta(a_t)\|^2 \\ &= \operatorname{dist}(a_t,\mathcal{K})^2 - 2\eta\,\langle g_\eta(a_t), a_t - x_t^\star\rangle + \eta^2\,\|g_\eta(a_t)\|^2 \\ &\le \operatorname{dist}(a_t,\mathcal{K})^2 - \eta^2\,\|g_\eta(a_t)\|^2 \\ &\le (1-\kappa^2)\operatorname{dist}(a_t,\mathcal{K})^2,\end{aligned}$$

where the last inequality follows from the Yosida error bound.

The same separation inequality and Cauchy–Schwarz imply

$$\eta\,\|g_\eta(a_t)\| \le \operatorname{dist}(a_t,\mathcal{K}).$$

Thus, if $\operatorname{dist}(a_t,\mathcal{K}) \le \rho_\eta(\varepsilon,R)$, then

$$s(a_t) = \eta\,\|g_\eta(a_t)\| \le \rho_\eta(\varepsilon,R).$$

By Lemma 23,

$$\operatorname{Gap}_\eta(a_t) \le \frac{s(a_t)^2 + 2Rs(a_t)}{\eta} \le \varepsilon.$$

Hence Lemma 6 certifies that $a_{t+1} = J_\eta(a_t)$ is an ε-SVI.

Since $\operatorname{dist}(a_0,\mathcal{K}) \le 2R$, the contraction established above yields

$$\operatorname{dist}(a_t,\mathcal{K}) \le 2R\exp(-\kappa^2 t/2).$$

The stated iteration bounds follow by requiring this distance to be at most $\rho_\eta(\varepsilon,R)$ and using $\rho_\eta(\varepsilon,R) = \Theta(\eta\varepsilon/R)$ when $\varepsilon \le R^2/\eta$.

Finally, let $u_{t+1} \in A(a_{t+1})$ be the returned SVI witness. By monotonicity, for every $y \in \mathcal{X}$ and $v \in A(y)$,

$$\langle v, y - a_{t+1}\rangle \ge \langle u_{t+1}, y - a_{t+1}\rangle \ge -\varepsilon,$$

which gives the ε-Minty guarantee. □

Metric subregularity of the constrained operator provides a sufficient condition for the Yosida error bound.

Proposition 49. *Assume A is monotone and let*

$$T := A + N_{\mathcal{X}}, \qquad Z := \{x \in \mathcal{X} : 0 \in T(x)\} \ne \emptyset.$$

Suppose T is globally metrically subregular for 0 on $\mathcal{X}$ with modulus $\lambda \ge 0$, meaning that

$$\operatorname{dist}(x,Z) \le \lambda\,\operatorname{dist}(0,T(x)) \qquad \text{for all } x \in \mathcal{X}.$$

Then, at any scale $\eta > 0$ with exact constrained-resolvent access, the Yosida error bound holds with

$$\kappa = \frac{\eta}{\eta+\lambda}.$$

Proof. Fix $a \in \mathcal{X}$ and let

$$x^+ := J_\eta(a), \qquad g_\eta(a) = \frac{a - x^+}{\eta}.$$

Since $g_\eta(a) \in T(x^+)$, metric subregularity gives

$$\operatorname{dist}(x^+,Z) \le \lambda\,\operatorname{dist}(0,T(x^+)) \le \lambda\,\|g_\eta(a)\|.$$

By the triangle inequality,

$$\begin{aligned}\operatorname{dist}(a,Z) &\le \|a - x^+\| + \operatorname{dist}(x^+,Z) \\ &\le (\eta+\lambda)\,\|g_\eta(a)\|.\end{aligned}$$

Since A is monotone, $Z \subseteq \mathcal{K}$. Consequently,

$$\operatorname{dist}(a,\mathcal{K}) \le \operatorname{dist}(a,Z) \le (\eta+\lambda)\,\|g_\eta(a)\|.$$

Rearranging yields

$$\eta\,\|g_\eta(a)\| \ge \frac{\eta}{\eta+\lambda}\operatorname{dist}(a,\mathcal{K}),$$

as claimed. □

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: Complete proofs are provided.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: Concretely in the main text.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

Justification: Complete proofs and clear assumptions are provided.

Guidelines:

- The answer [N/A] means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [N/A]

   Justification: [N/A]

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [N/A]

Justification: [N/A]

Guidelines:

- The answer [N/A] means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (https://neurips.cc/public/guides/CodeSubmissionPolicy) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (https://neurips.cc/public/guides/CodeSubmissionPolicy) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [N/A]

   Justification: [N/A]

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [N/A]

   Justification: [N/A]

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
   - The assumptions made should be given (e.g., Normally distributed errors).
   - It should be clear whether the error bar is the standard deviation or the standard error of the mean.

- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [N/A]

   Justification: [N/A]

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics https://neurips.cc/public/EthicsGuidelines?

   Answer: [Yes]

   Justification: This is primarily a theoretical study.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [No]

    Justification: This is primarily a theoretical study.

    Guidelines:

    - The answer [N/A] means that there is no societal impact of the work performed.
    - If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.

- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [No]

    Justification: This is primarily a theoretical study.

    Guidelines:

    - The answer [N/A] means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [No]

    Justification: This is primarily a theoretical study.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.

- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [N/A]

    Justification: [N/A]

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: This is primarily a theoretical study.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: This is primarily a theoretical study.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

Answer: [N/A]

Justification: [N/A]

Guidelines:

- The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.