# TOWARDS PRACTICAL GENERIC CONIC OPTIMIZATION[*]

CHRIS COEY[†], LEA KAPELEVICH[†], AND JUAN PABLO VIELMA[‡]

**Abstract.** Many convex optimization problems can be represented through conic *extended formulations* with auxiliary variables and constraints using only the small number of *standard cones* recognized by advanced conic solvers such as MOSEK 9. Such extended formulations are often significantly larger and more complex than equivalent conic *natural formulations*, which can use a much broader class of *exotic cones*. We define an exotic cone as a proper cone for which we can implement efficient logarithmically homogeneous self-concordant barrier oracles for the cone or its dual. Our goal is to establish whether a generic conic interior point method supporting natural formulations can outperform an advanced conic solver specialized for standard cones. We introduce Hypatia, a highly-configurable open-source conic primal-dual interior point solver with a generic interface for exotic cones. Hypatia is written in Julia and accessible through JuMP, and currently implements several dozen useful predefined cones. We define a subset of these cones, including some that have not been implemented before, and we propose several new efficient logarithmically homogeneous self-concordant barriers. We also describe and analyze techniques for constructing extended formulations for exotic conic constraints. For optimization problems from a variety of applications, we introduce natural formulations using our exotic cones, and we show that the natural formulations have much smaller dimensions and often lower barrier parameters than the equivalent extended formulations. Our computational experiments demonstrate the potential advantages, especially in terms of solve time and memory usage, of solving natural formulations with Hypatia compared to solving extended formulations with Hypatia or MOSEK.

**Key words.** conic optimization, extended formulations, interior point methods, logarithmically homogeneous self-concordant barrier functions

**AMS subject classifications.** 90-04, 90-08, 90C06, 90C22, 90C23, 90C25, 90C51

**1. Introduction.** Any convex optimization problem may be represented as a conic problem that minimizes a linear function over the intersection of an affine subspace with a Cartesian product of primitive proper cones (i.e. irreducible, closed, convex, pointed, and full-dimensional conic sets). An advantage of using conic form is that a conic problem, if well-posed, has a very simple and easily checkable certificate of optimality, primal infeasibility, or dual infeasibility.[1] We describe conic form and duality in section 3. Although advanced conic solvers currently recognize at most only a handful of *standard cones* (nonnegative, second order, rotated second order, positive semidefinite (PSD), and 3-dimensional exponential and power cones), these cones are sufficient for representing many problems of interest [21, 22]. Modeling tools such as disciplined convex programming (DCP) packages (see CVX [16], CVXPY [13], and Convex.jl [38]) and MathOptInterface's bridges [20] are designed to facilitate transformations of convex problems into conic problems with standard cones to enable access to powerful specialized conic solvers.

However, for many problems of interest, a representation in terms of standard cones is not the most natural or efficient conic representation. The process of transforming a general conic problem into a conic *extended formulation* (EF) that uses only

[†]Operations Research Center, MIT, Cambridge, MA (coey@mit.edu, lkap@mit.edu).

[‡]Google Research and MIT Sloan School of Management, Cambridge, MA (jvielma@google.com, jvielma@mit.edu).

[1]We limit the scope of this paper to conic problems, but note that there are other useful notions of duality that can be leveraged by convex optimization solvers; see for example DDS solver [19].

standard cones often requires introducing many auxiliary variables and constraints. If conic solvers could recognize a much larger class of cones, they could directly solve simpler, smaller conic *natural formulations* (NFs). This raises the question of whether it can be more efficient to solve NFs using a generic conic algorithm than to solve equivalent EFs using an advanced conic solver specialized for standard cones.

To answer this question, we develop a performant generic conic primal-dual interior point solver, Hypatia.[2] For various examples, we formulate NFs in section 6 using a variety of *exotic cones* defined in section 4. Several of the NFs and exotic cones have not been described or implemented before, and for our new cones we propose efficient logarithmically homogeneous self-concordant barrier (LHSCB) functions. In section 5 we describe general techniques for constructing EFs from NFs, and analyze some computational properties of these equivalent representations. For each of our example NFs, we randomly generate instances of a wide range of sizes. We construct equivalent EFs for the NF instances, and observe that the EFs have larger dimensions and often larger barrier parameter values. We demonstrate significant computational advantages from solving the NFs with Hypatia compared to solving the EFs with either Hypatia or the state-of-the-art specialized conic solver MOSEK 9.

**1.1. Conic primal-dual interior point methods.** Most successful commercial and open-source conic solvers (such as CSDP [8], CVXOPT [3], ECOS [36], MOSEK [22], SDPA [41]) implement primal-dual interior point methods (PDIPMs). Complexity analysis of PDIPMs, which relies on properties of LHSCBs, shows they require fewer iterations to converge but exhibit higher per-iteration cost compared to first order conic methods (see [30] on SCS solver). Computational evidence accords with this result and demonstrates the superior numerical robustness of PDIPMs.

Historically, PDIPM solvers were based on efficient algorithms specialized for symmetric cones, in particular, the nonnegative, (rotated) second order, and PSD cones. However, many useful non-symmetric conic constraints (such as $u \leq \log(w)$, representable with an exponential cone) are not representable with symmetric cones. Early non-symmetric conic PDIPMs such as [29, 26] had several disadvantages compared to the specialized symmetric methods, for example requiring a strictly feasible initial iterate, the solution of larger linear systems, and efficient oracles for LHSCBs of both primal and dual cones.

To address these issues, [37] introduced a PDIPM that requires only a few primal cone oracles: an initial interior point, feasibility check, and gradient and Hessian evaluations for an LHSCB. Starting from an initial iterate for the homogeneous self-dual embedding (HSDE) [2, 40], this algorithm approximately traces the central path through a series of iterations converging to a feasible solution for the HSDE, from which conic certificates may be obtained. Central path proximity is key to the polynomial time convergence analysis presented by [37] and later revised by [32]. After [37] demonstrated the practicality of their method on example formulations with non-symmetric 3-dimensional exponential and power cones, several conic solvers, including MOSEK 9 and the MATLAB solver Alfonso [33], implemented support for these cones.

**1.2. Natural and extended formulations.** Constructing an EF from an NF with exotic cones often requires introducing many artificial variables, linear equalities, or higher-dimensional conic constraints.[3] For example, in our density estimation

---

[2]Hypatia is available at github.com/chriscoey/Hypatia.jl under the permissive MIT license.

[3]EFs can be beneficial for accelerating outer approximation algorithms for mixed-integer conic optimization, such as the method implemented in Pajarito solver [11]. However, folklore says that

example problem in subsection 6.7, the dimensions of the variables, equalities, and conic constraints are typically orders of magnitude larger for the EFs than for the NFs. By increasing the size of problem data, EFs require larger linear systems to be solved throughout a PDIPM, worsening the per-iteration time and memory bottleneck. EFs are often associated with larger values of the barrier parameter $\nu$, which impacts the number of iterations $\mathcal{O}(\sqrt{\nu} \log \varepsilon^{-1})$ needed in the worst case to obtain a solution within $\varepsilon$ tolerance [28]. For example, in our matrix completion example problem in subsection 6.2, the NF uses a spectral norm cone with parameter $1 + d_1$ and the EF uses a PSD cone with parameter $d_1 + d_2$, where $d_1 \leq d_2$ are matrix side dimensions.

In section 5 we describe general techniques, some new, for constructing EFs from NFs. We analyze how these EF techniques necessarily increase the dimensions and often the barrier parameters associated with equivalent formulations. There may be other practical reasons to prefer NFs. Since EFs are often larger and more complex than NFs, they can be much slower and more memory-intensive to construct using modeling software, and less convenient for the modeler. Converting conic certificates from the space of the EF back into the more meaningful NF space can be complicated. Furthermore, the convergence conditions used by PDIPMs can provide numerical guarantees about conic certificates, but if EF certificates are converted to NF space, the NF certificates might lack such guarantees.

The potential computational advantage of NFs in the particular context of polynomial weighted sum-of-squares (SOS) optimization is illustrated by [33]. The authors describe LHSCBs for dual SOS cones, noting that efficient LHSCBs are not available for primal SOS cones (see subsection 4.12). They formulate dual SOS NFs directly and construct EFs with PSD cones. Although these EFs have the same barrier parameters as the NFs, they tend to be much larger. After implementing the basic PDIPM from [37, 32] in Alfonso, the authors observe improved solve times and scaleability from solving the SOS NFs with Alfonso compared to solving the EFs with MOSEK.

**1.3. A generic conic solver.** Our goal of broadening the computational argument for NFs motivates Hypatia's generic cone interface, which allows defining new primitive proper cones. The interface, like that of Alfonso, requires only the implementation of the few primal cone oracles needed by [37]. In section 4 we describe a subset of Hypatia's predefined cones and LHSCBs. Our descriptions, LHSCBs, and implementations for the logarithm cone, sparse PSD cone (with general non-chordal sparsity), root-determinant cone, log-determinant cone, and polynomial weighted SOS matrix cone in subsections 4.6, 4.9 to 4.11 and 4.13 are new, to our knowledge.

Thanks to an algorithmic innovation in Hypatia, defining a new cone through Hypatia's cone interface makes both the cone and its dual cone available for use in conic formulations. Since for many cones of interest, useful LHSCBs are only known for either the primal cone or the dual cone but not both (for example, a primal LHSCB for the spectral norm cone [17], and a dual LHSCB for the SOS cone [33]), Hypatia is able to solve a broader class of conic formulations than [37] and Alfonso, which can only handle cones with efficient primal oracles. For example, in our portfolio rebalancing example NF in subsection 6.1, we have $\ell_1$ norm cone and $\ell_\infty$ norm cone constraints; we are aware of an efficient LHSCB for the $\ell_\infty$ norm cone, but not for its dual cone, the $\ell_1$ norm cone (see subsection 4.2).

Hypatia is written in the Julia language [7] and is accessible through a powerful native interface or the convenient open-source modeling tool JuMP [15]. Hypatia

---

the EF for the second order cone likely slows down the conic solver, which is why Pajarito manages the EF in the MILP outer approximation model and only solves NFs for the conic subproblems.

has several notable algorithmic and software features that make it competitive and highly extensible. For example, although the PDIPM by [37] alternates between prediction and correction steps, Hypatia's default interior point algorithm uses a novel combined directions method with optional efficient third-order corrections, inspired by techniques from performant PDIPMs such as [3, 12, 14]. We defer a more detailed description of Hypatia's interfaces and algorithmic components to future work.

**1.4. Computational comparisons.** In section 6 we present a series of example problems from applications such as matrix completion, experiment design, and smooth density optimization. For these examples, we describe simple compact NFs using the cones we define in section 4. Some of these NFs are new and may be valuable to try in real-world applications. For each example NF, we randomly generate instances of a wide variety of sizes, and for each instance we empirically compare the dimensions and barrier parameters associated with the NF and its equivalent EF constructed using the techniques in section 5. Compared to the NFs, the EFs have significantly larger dimensions and equal or larger parameters, and are typically slower and more memory-intensive to construct using JuMP. We demonstrate significant improvements in solve time and memory usage from solving the NFs with Hypatia compared to solving the EFs with Hypatia or MOSEK.

**2. Notation.** For sets, cl denotes the closure and int denotes the interior. $\mathbb{R}$ is the scalar reals, $\mathbb{R}_{\geq}$ is the nonnegative reals, and $\mathbb{R}_{>} = \mathrm{int}(\mathbb{R}_{\geq})$ is the positive reals, $\mathbb{R}_{\leq}$ is the nonpositive reals, and $\mathbb{R}_{<} = \mathrm{int}(\mathbb{R}_{\leq})$ is the negative reals. The set of $d$-dimensional real vectors is $\mathbb{R}^d$, and the set of $d_1$-by-$d_2$-dimensional real matrices is $\mathbb{R}^{d_1 \times d_2}$. $\mathbb{S}^d$ is the set of symmetric matrices of side dimension $d$, $\mathbb{S}^d_{\succeq} \subset \mathbb{S}^d$ is the positive semidefinite matrices, and $\mathbb{S}^d_{\succ} = \mathrm{int}(\mathbb{S}^d_{\succeq})$ is the positive definite matrices. For some natural number $d$, $[\![d]\!]$ is the index set $\{1, 2, \ldots, d\}$.

If $a, b, c, d$ are scalars, vectors, or matrices (of appropriate dimensions), the notation $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ usually denotes concatenation into a matrix. For a vector or matrix $A$, the transpose is $A'$. $I(d)$ is the identity matrix in $\mathbb{R}^{d \times d}$. For dimensions implied by context, 0 may represent vectors or matrices of 0s, and $e$ is a vector of 1s. Diag represents the diagonal matrix of a given vector, and diag represents the diagonal vector of a given square matrix. The inner product of vectors $u, w \in \mathbb{R}^d$ is $u'w = \sum_{i \in [\![d]\!]} u_i w_i$. log is the natural logarithm, $\|\cdot\|_p$ is the $\ell_p$ norm (for $p \geq 1$) of a vector, det is the determinant of a symmetric matrix, tr is the matrix trace, and $\sigma_i(\cdot)$ is the $i$th largest singular value of a matrix.

For a function $f : \mathbb{R}^d \to \mathbb{R}$ and a point $p \in \mathbb{R}^d$, the Hessian (of second order partial derivatives) of $f$ at $p$ is $\nabla^2 f(p) \in \mathbb{S}^d$. Given a direction $h \in \mathbb{R}^d$, the second and third order directional derivatives of $f$ at $p$ are $\nabla^2 f(p)[h, h] \in \mathbb{R}$ and $\nabla^3 f(p)[h, h, h] \in \mathbb{R}$.

The operator vec maps $\mathbb{R}^{d_1 \times d_2}$ (matrices) to $\mathbb{R}^{d_1 d_2}$ (vectors) by stacking columns. The inverse operator $\mathrm{mat}_{d_1, d_2}$ maps $\mathbb{R}^{d_1 d_2}$ to $\mathbb{R}^{d_1 \times d_2}$. For symmetric matrices, vec maps $\mathbb{S}^d$ to $\mathbb{R}^{\mathrm{sd}(d)}$, where $\mathrm{sd}(d) = d(d+1)/2$, by rescaling off-diagonal elements by $\sqrt{2}$ and stacking columns of the upper triangle. For example, for $S \in \mathbb{S}^3$ we have $\mathrm{sd}(3) = 6$ and $\mathrm{vec}(S) = (S_{1,1}, \sqrt{2}S_{1,2}, S_{2,2}, \sqrt{2}S_{1,3}, \sqrt{2}S_{2,3}, S_{3,3}) \in \mathbb{R}^{\mathrm{sd}(3)}$. The inverse mapping mat from $\mathbb{R}^{\mathrm{sd}(d)}$ to $\mathbb{S}^d$ is well-defined. The linear operators vec and mat preserve inner products, so $\mathrm{vec}(S)' \mathrm{vec}(Z) = \mathrm{tr}(S'Z)$ for $S, Z \in \mathbb{R}^{d_1 \times d_2}$ or $S, Z \in \mathbb{S}^d$.

**3. Conic duality and standard form.** Let $\mathcal{K}$ be a proper cone in $\mathbb{R}^q$, i.e. a conic subset of $\mathbb{R}^q$ that is closed, convex, pointed, and full-dimensional (see [37]). We call $\mathcal{K}$ a primitive (or irreducible) cone if it cannot be written as a Cartesian product

of two or more lower-dimensional cones. $\mathcal{K}^* \subset \mathbb{R}^q$ is the dual cone of $\mathcal{K}$:

$$\text{(3.1)} \qquad \mathcal{K}^* = \{z \in \mathbb{R}^q : s'z \geq 0, \forall s \in \mathcal{K}\}.$$

$\mathcal{K}^*$ is a primitive proper cone if and only if $\mathcal{K}$ is a primitive proper cone.

Hypatia's convenient primal conic form over variable $x \in \mathbb{R}^n$ is:

$$\text{(3.2a)} \qquad \inf_x \quad c'x :$$
$$\text{(3.2b)} \qquad b - Ax = 0,$$
$$\text{(3.2c)} \qquad h - Gx \in \mathcal{K},$$

where $c \in \mathbb{R}^n$, $b \in \mathbb{R}^p$, and $h \in \mathbb{R}^q$ are vectors, $A : \mathbb{R}^n \to \mathbb{R}^p$ and $G : \mathbb{R}^n \to \mathbb{R}^q$ are linear maps, and $\mathcal{K} \subset \mathbb{R}^q$ is a Cartesian product $\mathcal{K} = \mathcal{K}_1 \times \cdots \times \mathcal{K}_K$ of primitive proper cones. Henceforth we use $n, p, q$ to denote the variable, equality, and conic constraint dimensions of a conic problem. The corresponding conic dual problem over variable $y \in \mathbb{R}^p$ associated with (3.2b), and $z \in \mathbb{R}^q$ associated with (3.2c), is:

$$\text{(3.3a)} \qquad \sup_{y,z} \quad -b'y - h'z :$$
$$\text{(3.3b)} \qquad c + A'y + G'z = 0,$$
$$\text{(3.3c)} \qquad z \in \mathcal{K}^*,$$

where (3.3b) is associated with primal variable $x \in \mathbb{R}^n$. Note $\mathcal{K}^* = \mathcal{K}_1^* \times \cdots \times \mathcal{K}_K^*$.

If the conic primal-dual pair (3.2)–(3.3) is well-posed, there exist simple conic certificates providing easily verifiable proofs of infeasibility of the primal or dual problems or optimality of a given primal-dual solution [39, 34]. A primal improving ray $x$ is a feasible direction for the primal along which the primal objective improves (i.e. $c'x < 0$, $-Ax = 0$, $-Gx \in \mathcal{K}$), and hence it certifies dual infeasibility via the conic generalization of Farkas' lemma. Similarly, a dual improving ray $(y, z)$ certifies primal infeasibility (i.e. $-b'y - h'z > 0$, $A'y + G'z = 0$, $z \in \mathcal{K}^*$). Finally, a complementary solution $(x, y, z)$ satisfies the primal-dual feasibility conditions (3.2b)–(3.2c) and (3.3b)–(3.3c) and has equal and attained objective values $c'x = -b'y - h'z$, and hence certifies optimality of $(x, y, z)$ via conic weak duality.

**4. Predefined primitive proper cones.** Hypatia's generic cone interface allows defining any primitive proper cone $\mathcal{K}$ by specifying a small list of oracles: an initial interior point $t \in \text{int}(\mathcal{K})$, a feasibility check for $\text{int}(\mathcal{K})$, and gradients and Hessians of an LHSCB $f$ for $\mathcal{K}$. Following [27, sections 2.3.1 and 2.3.3], a three times continuously differentiable convex function $f$, defined on $\text{int}(\mathcal{K})$, is an LHSCB for $\mathcal{K} \subset \mathbb{R}^q$ if $f(w_i) \to \infty$ along every sequence $w_i \in \text{int}(\mathcal{K})$ converging to the boundary of $\mathcal{K}$, and:

$$\text{(4.1a)} \qquad |\nabla^3 f(w)[h, h, h]| \leq 2|\nabla^2 f(w)[h, h]|^{3/2} \qquad \forall w \in \mathcal{K}, h \in \mathbb{R}^q,$$
$$\text{(4.1b)} \qquad f(\theta w) = f(w) - \nu \log(\theta) \qquad \forall w \in \mathcal{K}, \theta \in \mathbb{R},$$

where $\nu \in \mathbb{R}$ in (4.1b) is the barrier parameter of $f$ (as we note in subsection 1.2, $\nu$ impacts the worst case iteration complexity for PDIPMs). The interface also allows optional specification of other oracles that can improve performance. Once defined, the cone and its dual cone may be used in any combination with other cones recognized by Hypatia to construct the Cartesian product cone in (3.2c).

In subsections 4.1 to 4.13, we introduce a subset of Hypatia's predefined primitive proper cones, their dual cones, and associated LHSCBs. We use these cones in our

5

example NFs and EFs in section 6. The techniques and references in this section may provide helpful examples for readers trying to develop new cones and LHSCBs. Several of our cone definitions and LHSCBs appear to be novel, in particular, our logarithm cone, sparse PSD cone (with general non-chordal sparsity), and polynomial weighted SOS matrix cone in subsections 4.6, 4.9 and 4.13, which rely on [27, Proposition 5.1.1], and our root-determinant and log-determinant cones in subsections 4.10 and 4.11, which rely on [27, Definitions 2.3.2 and 5.1.1 and Proposition 5.1.7].

**4.1. Nonnegative cone.** The self-dual nonnegative cone is $\mathcal{K}_{\mathbb{R}} = \mathcal{K}_{\mathbb{R}}^* = \mathbb{R}_{\geq}$. We use the LHSCB $f(w) = -\log(w)$ from [28, section 2.1] with $\nu = 1$.

**4.2. Infinity norm cone.** The $\ell_\infty$ norm cone is the epigraph of $\ell_\infty$, and its dual cone is the $\ell_1$ norm cone:

$$(4.2a) \qquad \mathcal{K}_{\ell_\infty} = \{(u, w) \in \mathbb{R}_{\geq} \times \mathbb{R}^d : u \geq \|w\|_\infty\},$$

$$(4.2b) \qquad \mathcal{K}_{\ell_\infty}^* = \{(u, w) \in \mathbb{R}_{\geq} \times \mathbb{R}^d : u \geq \|w\|_1\}.$$

We are not aware of a useful LHSCB for $\mathcal{K}_{\ell_\infty}^*$, but for $\mathcal{K}_{\ell_\infty}$ we use the LHSCB from [17, section 7.5] with $\nu = 1 + d$:

$$(4.3) \qquad f(u, w) = (d-1)\log(u) - \sum_{i \in \llbracket d \rrbracket} \log(u^2 - w_i^2).$$

**4.3. Euclidean norm cone.** The self-dual Euclidean norm cone (or second order cone) is the epigraph of the $\ell_2$ norm:

$$(4.4) \qquad \mathcal{K}_{\ell_2} = \mathcal{K}_{\ell_2}^* = \{(u, w) \in \mathbb{R}_{\geq} \times \mathbb{R}^d : u \geq \|w\|_2\}.$$

We use the LHSCB from [28, section 2.3] with $\nu = 2$:

$$(4.5) \qquad f(u, w) = -\log(u^2 - \|w\|_2^2).$$

**4.4. Euclidean norm-squared cone.** The self-dual Euclidean norm-squared cone (or rotated second order cone) is the epigraph of the perspective of $g(w) = \frac{1}{2}\|w\|_2^2$:

$$(4.6) \qquad \mathcal{K}_{\ell_2^2} = \mathcal{K}_{\ell_2^2}^* = \{(u, v, w) \in \mathbb{R}_{\geq}^2 \times \mathbb{R}^d : 2uv \geq \|w\|_2^2\}.$$

We use the LHSCB from [28, section 2.3] with $\nu = 2$:

$$(4.7) \qquad f(u, v, w) = -\log(2uv - \|w\|_2^2).$$

**4.5. Geometric mean cone.** The geometric mean cone and its dual cone are:

$$(4.8a) \qquad \mathcal{K}_{\text{geo}} = \{(u, w) \in \mathbb{R} \times \mathbb{R}_{\geq}^d : u \leq \prod_{i \in \llbracket d \rrbracket} w_i^{1/d}\},$$

$$(4.8b) \qquad \mathcal{K}_{\text{geo}}^* = \{(u, w) \in \mathbb{R}_{\leq} \times \mathbb{R}_{\geq}^d : \frac{-u}{d} \leq \prod_{i \in \llbracket d \rrbracket} w_i^{1/d}\}.$$

We use the LHSCB from [25, section 4] with $\nu = 1 + d$:

$$(4.9) \qquad f(u, w) = -\sum_{i \in \llbracket d \rrbracket} \log(w_i) - \log(\prod_{i \in \llbracket d \rrbracket} w_i^{1/d} - u).$$

**4.6. Logarithm cone.** The logarithm cone is the hypograph of the perspective of a sum of log functions:

$$(4.10a) \qquad \mathcal{K}_{\log} = \text{cl}\{(u, v, w) \in \mathbb{R} \times \mathbb{R}_{>}^{1+d} : u \leq \sum_{i \in \llbracket d \rrbracket} v \log(\tfrac{w_i}{v})\},$$

$$(4.10b) \qquad \mathcal{K}_{\log}^* = \text{cl}\{(u, v, w) \in \mathbb{R}_{<} \times \mathbb{R} \times \mathbb{R}_{>}^d : v \geq \sum_{i \in \llbracket d \rrbracket} u(\log(\tfrac{-w_i}{u}) + 1)\}.$$

We propose an LHSCB in Lemma A.1 with $\nu = 2 + d$:

$$(4.11) \qquad f(u, v, w) = -\log(v) - \sum_{i \in \llbracket d \rrbracket} \log(w_i) - \log(\sum_{i \in \llbracket d \rrbracket} v \log(\tfrac{w_i}{v}) - u).$$

**4.7. Spectral norm cone.** The spectral norm cone is the epigraph of the matrix spectral norm, and its dual cone is the epigraph of the matrix nuclear norm:

$$(4.12a) \qquad \mathcal{K}_{\mathrm{spec}(r,s)} = \{(u,w) \in \mathbb{R}_\geq \times \mathbb{R}^{rs} : u \geq \sigma_1(W)\},$$

$$(4.12b) \qquad \mathcal{K}^*_{\mathrm{spec}(r,s)} = \{(u,w) \in \mathbb{R}_\geq \times \mathbb{R}^{rs} : u \geq \sum_{i \in [\![r]\!]} \sigma_i(W)\},$$

where $W = \mathrm{mat}_{r,s}(w) \in \mathbb{R}^{r \times s}$ and $r \leq s$ (this is nonrestrictive since the singular values are the same for $W$ and $W'$). We are not aware of a useful LHSCB for $\mathcal{K}^*_{\mathrm{spec}}$, but for $\mathcal{K}_{\mathrm{spec}}$ we use the LHSCB from [27] with $\nu = 1 + r$:

$$(4.13) \qquad f(u,w) = -\log(u) - \mathrm{logdet}(uI(r) - \tfrac{WW'}{u}).$$

**4.8. Positive semidefinite cone.** The self-dual positive semidefinite cone is:

$$(4.14) \qquad \mathcal{K}_\mathbb{S} = \mathcal{K}^*_\mathbb{S} = \{w \in \mathbb{R}^{\mathrm{sd}(d)} : \mathrm{mat}(w) \in \mathbb{S}^d_\succeq\}.$$

We use the LHSCB from [28, section 2.2] with $\nu = d$:

$$(4.15) \qquad f(w) = -\mathrm{logdet}(\mathrm{mat}(w)).$$

**4.9. Sparse positive semidefinite cone.** Suppose $\mathcal{S} = ((i_l, j_l))_{l \in [\![d]\!]}$ is a collection of row-column index pairs defining the sparsity pattern (including all diagonal elements) of the lower triangle of a symmetric matrix of side dimension $s$. Unlike prior work such as [10, 5], we do not require that $\mathcal{S}$ be a chordal sparsity pattern, hence the cone dimension $d = |\mathcal{S}|$ can be as small as possible. Note $s \leq d \leq \mathrm{sd}(s)$. Let $\mathrm{mat}_\mathcal{S} : \mathbb{R}^d \to \mathbb{S}^s$ be the linear operator satisfying:

$$(4.16) \qquad (\mathrm{mat}_\mathcal{S}(w))_{i,j} = \begin{cases} w_l & \text{if } i = i_l = j = j_l, \\ \frac{w_l}{\sqrt{2}} & \text{if } i = i_l \neq j = j_l, \\ 0 & \text{otherwise,} \end{cases} \qquad \forall i,j \in [\![s]\!] : i \geq j.$$

We define the sparse PSD cone and its dual cone of PSD-completable matrices as:

$$(4.17a) \qquad \mathcal{K}_{\mathrm{spsd}(\mathcal{S})} = \{w \in \mathbb{R}^d : \mathrm{mat}_\mathcal{S}(w) \in \mathbb{S}^s_\succeq\},$$

$$(4.17b) \qquad \mathcal{K}^*_{\mathrm{spsd}(\mathcal{S})} = \{w \in \mathbb{R}^d : \exists \theta \in \mathbb{R}^{\mathrm{sd}(s)-d}, \mathrm{mat}_\mathcal{S}(w) + \mathrm{mat}_{\bar{\mathcal{S}}}(\theta) \in \mathbb{S}^s_\succeq\},$$

where $\bar{\mathcal{S}}$ is the lower triangle inverse sparsity pattern of $\mathcal{S}$ (with $|\bar{\mathcal{S}}| = \mathrm{sd}(s) - d$). Noting that (4.17a) implicitly constrains a linear function of $w$ to a PSD cone, using the $\mathcal{K}_\mathbb{S}$ barrier (4.15) with [27, Proposition 5.1.1], we have the LHSCB with $\nu = s$:

$$(4.18) \qquad f(w) = -\mathrm{logdet}(\mathrm{mat}_\mathcal{S}(w)).$$

**4.10. Root-determinant cone.** The root-determinant cone is the hypograph of the root-determinant function:

$$(4.19a) \qquad \mathcal{K}_{\mathrm{rtdet}} = \{(u,w) \in \mathbb{R}^{1+\mathrm{sd}(d)} : W \in \mathbb{S}^d_\succeq, u^d \leq \det(W)\},$$

$$(4.19b) \qquad \mathcal{K}^*_{\mathrm{rtdet}} = \{(u,w) \in \mathbb{R}_\leq \times \mathbb{R}^{\mathrm{sd}(d)} : W \in \mathbb{S}^d_\succeq, (\tfrac{-u}{d})^d \leq \det(W)\},$$

where $W = \mathrm{mat}(w)$. We use the LHSCB suggested by A. Nemirovski (see Proposition A.3) with $\nu = (\tfrac{5}{3})^2(1+d)$:

$$(4.20) \qquad f(u,w) = -(\tfrac{5}{3})^2(\mathrm{logdet}(W) + \log((\det(W))^{1/d} - u)).$$

**4.11. Log-determinant cone.** The log-determinant cone is the hypograph of the perspective of the log-determinant function:

(4.21a)   $\mathcal{K}_{\text{logdet}} = \text{cl}\{(u, v, w) \in \mathbb{R} \times \mathbb{R}_> \times \mathbb{R}^{\text{sd}(d)} : W \in \mathbb{S}_{\succ}^d, u \leq v \, \text{logdet}(\frac{W}{v})\}$,

(4.21b)   $\mathcal{K}_{\text{logdet}}^* = \text{cl}\{(u, v, w) \in \mathbb{R}_< \times \mathbb{R}^{1+\text{sd}(d)} : W \in \mathbb{S}_{\succ}^d, v \geq u(\text{logdet}(\frac{-W}{u}) + d)\}$,

where $W = \text{mat}(w)$. We propose an LHSCB in Proposition A.2 with $\nu = 2 + d$:

(4.22)   $f(u, v, w) = -\log(v) - \text{logdet}(W) - \log(v \, \text{logdet}(\frac{W}{v}) - u)$.

**4.12. Polynomial weighted sum-of-squares cone.** Given a collection of matrices $P_l \in \mathbb{R}^{d \times t_l}, \forall l \in [\![r]\!]$ derived from basis polynomials evaluated at $d$ interpolation points as in [33], the interpolant basis polynomial weighted SOS cone is:

(4.23a)   $\mathcal{K}_{\text{sos}(P)} = \{w \in \mathbb{R}^d : \exists \Theta_l \in \mathbb{S}_{\succeq}^{t_l}, \forall l \in [\![r]\!], w = \sum_{l \in [\![r]\!]} \text{diag}(P_l \Theta_l P_l')\}$,

(4.23b)   $\mathcal{K}_{\text{sos}(P)}^* = \{w \in \mathbb{R}^d : P_l' \, \text{Diag}(w) P_l \in \mathbb{S}_{\succeq}^{t_l}, \forall l \in [\![r]\!]\}$.

These cones are useful for polynomial and moment modeling; for example, a point in $\mathcal{K}_{\text{sos}(P)}$ specifies a polynomial that is pointwise nonnegative on a semialgebraic domain defined by $P$. According to [33], a useful LHSCB is not known for $\mathcal{K}_{\text{sos}(P)}$, but an LHSCB for $\mathcal{K}_{\text{sos}(P)}^*$ with $\nu = \sum_{l \in [\![r]\!]} t_l$ is:

(4.24)   $f(w) = -\sum_{l \in [\![r]\!]} \text{logdet}(P_l' \, \text{Diag}(w) P_l)$.

**4.13. Polynomial weighted sum-of-squares matrix cone.** Given a side dimension $s$ of a symmetric matrix of polynomials (for simplicity, all using the same interpolant basis), and $P_l \in \mathbb{R}^{d \times t_l}, \forall l \in [\![r]\!]$ defined as for $\mathcal{K}_{\text{sos}(P)}$ in subsection 4.12, the interpolant basis polynomial weighted SOS matrix cone is:

(4.25a)   $\mathcal{K}_{\text{sosm}(P)} = \left\{ \begin{array}{l} w \in \mathbb{R}^{\text{sd}(s)d} : \exists \Theta_l \in \mathbb{S}_{\succeq}^{st_l}, \forall l \in [\![r]\!], \\ W_{i,j,:} = \sum_{l \in [\![r]\!]} \text{diag}(P_l(\Theta_l)_{i,j} P_l'), \forall i, j \in [\![s]\!] : i \geq j \end{array} \right\}$,

(4.25b)   $\mathcal{K}_{\text{sosm}(P)}^* = \{w \in \mathbb{R}^{\text{sd}(s)d} : [P_l' \, \text{Diag}(W_{i,j,:}) P_l]_{i,j \in [\![s]\!]} \in \mathbb{S}_{\succeq}^{st_l}, \forall l \in [\![r]\!]\}$,

where $W_{i,j,:} \in \mathbb{R}^d$ is the contiguous slice of $w$ (scaled to account for symmetry of the polynomial matrix) corresponding to the interpolant basis coefficients for the polynomial in the $(i, j)$th position of the polynomial matrix, $(S)_{i,j}$ is the $(i, j)$th block in a symmetric matrix $S$ with square blocks of equal dimensions, and $[g(W_{i,j,:})]_{i,j \in [\![s]\!]}$ is the symmetric matrix with square matrix $g(W_{i,j,:})$ in the $(i, j)$th block. A point in $\mathcal{K}_{\text{sosm}(P)}$ specifies a polynomial matrix that is pointwise PSD on a semialgebraic domain defined by $P$. We are not aware of a useful LHSCB for $\mathcal{K}_{\text{sosm}(P)}$ (indeed, for $s = 1$, $\mathcal{K}_{\text{sosm}(P)}$ reduces to $\mathcal{K}_{\text{sos}(P)}$). Noting that (4.25b) implicitly constrains a linear function of $w$ to a Cartesian product of PSD cones, using the $\mathcal{K}_{\mathbb{S}}$ barrier (4.15) with [27, Propositions 5.1.1 and 5.1.3] we propose the LHSCB for $\mathcal{K}_{\text{sosm}(P)}^*$ with $\nu = s \sum_{l \in [\![r]\!]} t_l$:

(4.26)   $f(w) = -\sum_{l \in [\![r]\!]} \text{logdet}([P_l' \, \text{Diag}(W_{i,j,:}) P_l]_{i,j \in [\![s]\!]})$.

**5. Natural and extended formulations for conic constraints.** In subsections 5.1 to 5.7, we describe general techniques for constructing standard cone EFs for the types of exotic conic constraints we use in our example NFs in section 6. Recall

that we define the standard cones as those recognized by MOSEK 9: $\mathcal{K}_{\mathbb{R}}$, $\mathcal{K}_{\ell_2}$, $\mathcal{K}_{\ell_2^2}$, $\mathcal{K}_{\mathbb{S}}$, and the 3-dimensional exponential and power cones. The exponential cone is a special case of our logarithm cone $\mathcal{K}_{\log}$ (let $d = 1$ in (4.10a)), so any 3-dimensional $\mathcal{K}_{\log}$ constraint is an exponential cone constraint. Although MOSEK's power cone in $\mathbb{R}^3$ is a special case of Hypatia's general power cone, we do not define power cones in section 4 because they are not useful for our NFs and EFs in section 6.

In this section and in section 6, we often refer to an exotic cone constraint as an NF and an equivalent reformulation of an NF constraint in terms of standard cones as an EF. In general, an NF constraint has the form $h - Gx \in \mathcal{K}$, but for convenience here we write $w \in \mathcal{K}$, since $w = h - Gx$ can be substituted into the EF description. An EF may use auxiliary variables (increasing the dimension $n$ of variable $x$ in the primal conic form (3.2)), equalities (increasing the equality constraint dimension $p$ in (3.2b)), and conic constraints (increasing the conic constraint dimension $q$ in (3.2c)). In Table 5.1, we compare the dimensions and barrier parameters associated with equivalent NF and EF constraints.

Some EFs here are new and others follow best practices from DCP modeling tools such as Convex.jl [38] and descriptions such as [6, chapter 4]. Since we use JuMP [15] to build NFs and EFs in section 6, we contributed several exotic cones and the EFs described in subsections 5.1, 5.2 and 5.4 to MathOptInterface's bridges [20], to permit automated EF construction. MathOptInterface does not currently recognize $\mathcal{K}_{\log}$ (for $d > 1$), $\mathcal{K}_{\text{spsd}}$, $\mathcal{K}_{\text{sos}}$, or $\mathcal{K}_{\text{sosm}}$, so we construct the EFs in subsections 5.3, 5.5 and 5.7 manually. For some EFs with auxiliary variables and equalities, it is possible to perform eliminations to reduce dimensions somewhat, but this can impact the sparsity of problem data (and in our experiments in section 6, both Hypatia and MOSEK perform preprocessing).

| NF | $q$ | $\nu$ | EF | $\bar{q}$ | $\bar{\nu}$ | $\bar{n}$ | $\bar{p}$ |
|---|---|---|---|---|---|---|---|
| $\mathcal{K}_{\ell_\infty}$ | $1+d$ | $1+d$ | $\mathcal{K}_{\mathbb{R}}$ | $2d$ | $2d$ | $0$ | $0$ |
| $\mathcal{K}_{\ell_\infty}^*$ | $1+d$ | $1+d$ | $\mathcal{K}_{\mathbb{R}}$ | $1+2d$ | $1+2d$ | $2d$ | $d$ |
| $\mathcal{K}_{\text{geo}}$ | $1+d$ | $1+d$ | $\mathcal{K}_{\mathbb{R}},\mathcal{K}_{\log}$ | $2+3d$ | $2+3d$ | $1+d$ | $0$ |
| $\mathcal{K}_{\log}$ | $2+d$ | $2+d$ | $\mathcal{K}_{\mathbb{R}},\mathcal{K}_{\log}$ | $1+3d$ | $1+3d$ | $d$ | $0$ |
| $\mathcal{K}_{\text{spec}}$ | $1+rs$ | $1+r$ | $\mathcal{K}_{\mathbb{S}}$ | $\text{sd}(r+s)$ | $r+s$ | $0$ | $0$ |
| $\mathcal{K}_{\text{spec}}^*$ | $1+rs$ | $1+r$ | $\mathcal{K}_{\mathbb{R}},\mathcal{K}_{\mathbb{S}}$ | $1+\text{sd}(r+s)$ | $1+r+s$ | $\text{sd}(r)+\text{sd}(s)$ | $0$ |
| $\mathcal{K}_{\text{spsd}}$ | $d$ | $s$ | $\mathcal{K}_{\mathbb{S}}$ | $\text{sd}(s)$ | $\nu$ | $0$ | $0$ |
| $\mathcal{K}_{\text{spsd}}^*$ | $d$ | $s$ | $\mathcal{K}_{\mathbb{S}}$ | $\text{sd}(s)$ | $\nu$ | $\text{sd}(s)-d$ | $0$ |
| $\mathcal{K}_{\text{rtdet}}$ | $1+\text{sd}(d)$ | $\frac{25}{9}(1+d)$ | $\mathcal{K}_{\mathbb{R}},\mathcal{K}_{\log},\mathcal{K}_{\mathbb{S}}$ | $2+3d+\text{sd}(2d)$ | $2+5d$ | $1+d+\text{sd}(d)$ | $0$ |
| $\mathcal{K}_{\text{logdet}}$ | $2+\text{sd}(d)$ | $2+d$ | $\mathcal{K}_{\mathbb{R}},\mathcal{K}_{\log},\mathcal{K}_{\mathbb{S}}$ | $1+3d+\text{sd}(2d)$ | $1+5d$ | $1+d+\text{sd}(d)$ | $0$ |
| $\mathcal{K}_{\text{sos}}$ | $d$ | $\sum_l t_l$ | $\mathcal{K}_{\mathbb{S}}$ | $\sum_l \text{sd}(t_l)$ | $\nu$ | $\bar{q}$ | $d$ |
| $\mathcal{K}_{\text{sos}}^*$ | $d$ | $\sum_l t_l$ | $\mathcal{K}_{\mathbb{S}}$ | $\sum_l \text{sd}(t_l)$ | $\nu$ | $0$ | $0$ |
| $\mathcal{K}_{\text{sosm}}$ | $\text{sd}(s)d$ | $s\sum_l t_l$ | $\mathcal{K}_{\mathbb{S}}$ | $\sum_l \text{sd}(st_l)$ | $\nu$ | $\bar{q}$ | $q$ |

Table 5.1: Computational properties of NFs and EFs of exotic conic constraints in subsections 5.1 to 5.7. $q$ and $\nu$ are the dimension and parameter for the NF cone, and $\bar{q}$ and $\bar{\nu}$ are the corresponding values for the EF Cartesian product cone. $\bar{n}$ and $\bar{p}$ are the auxiliary variable and equality dimensions for the EF. Note $\text{sd}(k)$ is $\mathcal{O}(k^2)$.

**5.1. Infinity norm cone.** Our examples in subsections 6.1 and 6.5 use the following NF (left) and EF (right):

(5.1) $$(u, w) \in \mathcal{K}_{\ell_\infty} \subset \mathbb{R}^{1+d} \Leftrightarrow (ue - w, ue + w) \in (\mathcal{K}_{\mathbb{R}})^{2d},$$

and similarly, subsection 6.1 uses:

$$(5.2) \qquad (u, w) \in \mathcal{K}_{\ell_\infty}^* \subset \mathbb{R}^{1+d} \Leftrightarrow \exists (\theta, \lambda) \in (\mathcal{K}_\mathbb{R})^{2d}, w = \theta - \lambda, u - e'(\theta + \lambda) \in \mathcal{K}_\mathbb{R}.$$

**5.2. Geometric mean cone.** The example in subsection 6.2 uses an EF for $\mathcal{K}_{\text{geo}}$, and the root-determinant variant of the example in subsection 6.5 uses an EF for $\mathcal{K}_{\text{geo}}$ indirectly through a $\mathcal{K}_{\text{rtdet}}$ EF (see subsection 5.6). We are aware of three EFs for $\mathcal{K}_{\text{geo}}$: a rotated second order cone EF (*EF-sec*) from [6, section 3.3.1], a power cone EF (*EF-pow*) from [22], and an exponential cone EF (*EF-exp*). We contributed EF-exp to MathOptInterface as a combination of two bridges (geometric mean cone to relative entropy cone to exponential cones):

$$(5.3) \qquad (u, w) \in \mathcal{K}_{\text{geo}} \subset \mathbb{R}^{1+d} \Leftrightarrow \begin{array}{l} \exists (\theta, \lambda) \in \mathbb{R}^{1+d}, (\theta, -e'\lambda) \in (\mathcal{K}_\mathbb{R})^2, \\ (-\lambda_i, u + \theta, w_i) \in \mathcal{K}_{\log}, \forall i \in [\![d]\!]. \end{array}$$

EF-pow is not currently available through MathOptInterface bridges, and it has a very similar size and structure to EF-exp, so we do not describe or test it. EF-sec uses multiple levels of variables and 3-dimensional $\mathcal{K}_{\ell_2^2}$ constraints and is complex to describe, so we refer the reader to [6, section 3.3.1]. In our empirical comparisons in subsections 6.2 and 6.5, EF-sec typically has larger variable and conic constraint dimensions but smaller barrier parameter than EF-exp.

**5.3. Logarithm cone.** Subsection 6.7 uses the EF (when $d > 1$):

$$(5.4) \quad (u, v, w) \in \mathcal{K}_{\log} \subset \mathbb{R}^{2+d} \Leftrightarrow \exists \theta \in \mathbb{R}^d, e'\theta - u \in \mathcal{K}_\mathbb{R}, (\theta_i, 1, w_i) \in \mathcal{K}_{\log}, \forall i \in [\![d]\!].$$

**5.4. Spectral norm cone.** Subsection 6.2 uses the EF from [6, section 4.2]:

$$(5.5) \qquad (u, w) \in \mathcal{K}_{\text{spec}(r,s)} \subset \mathbb{R}^{1+rs} \Leftrightarrow \begin{bmatrix} uI(r) & W \\ W' & uI(s) \end{bmatrix} \in \mathbb{S}_{\succeq}^{r+s}.$$

Subsection 6.3 uses the EF from [35], with $\Theta = \text{mat}(\theta) \in \mathbb{S}^r, \Lambda = \text{mat}(\lambda) \in \mathbb{S}^s$:

$$(5.6) \qquad (u, w) \in \mathcal{K}_{\text{spec}(r,s)}^* \subset \mathbb{R}^{1+rs} \Leftrightarrow \begin{array}{l} \exists (\theta, \lambda) \in \mathbb{R}^{\text{sd}(r)+\text{sd}(s)}, \begin{bmatrix} \Theta & W \\ W' & \Lambda \end{bmatrix} \in \mathbb{S}_{\succeq}^{r+s}, \\ u - (\text{tr}(\Theta) + \text{tr}(\Lambda))/2 \in \mathcal{K}_\mathbb{R}. \end{array}$$

**5.5. Sparse PSD cone.** Subsection 6.4 uses the EFs implicit in the definitions of $\mathcal{K}_{\text{spsd}}$ and $\mathcal{K}_{\text{spsd}}^*$ in (4.17a) and (4.17b). Note that $\mathcal{K}_{\text{spsd}}^*$ requires auxiliary variables.

**5.6. Root-determinant and log-determinant cones.** For $\theta \in \mathbb{R}^{\text{sd}(d)}$, let $\Theta = \text{mat}(\theta) \in \mathbb{S}^d$, and $D(W, \Theta) = \begin{bmatrix} W & \Theta \\ \Theta' & \text{Diag}(\text{diag}(\Theta)) \end{bmatrix} \in \mathbb{S}^{2d}$. Subsection 6.5 uses the EF adapted from [6, section 4.2]:

$$(5.7) \quad (u, w) \in \mathcal{K}_{\text{rtdet}} \subset \mathbb{R}^{1+\text{sd}(d)} \Leftrightarrow \exists \theta \in \mathbb{R}^{\text{sd}(d)}, D(W, \Theta) \in \mathbb{S}_{\succeq}^{2d}, (u, \text{diag}(\Theta)) \in \mathcal{K}_{\text{geo}},$$

where the $\mathcal{K}_{\text{geo}}$ constraint must be replaced with one of the geometric mean cone EFs described subsection 5.2. Similarly, subsection 6.5 uses a related EF:

$$(5.8) \quad (u, v, w) \in \mathcal{K}_{\text{logdet}} \subset \mathbb{R}^{2+\text{sd}(d)} \Leftrightarrow \begin{array}{l} \exists (\theta, \lambda) \in \mathbb{R}^{\text{sd}(d)+d}, e'\lambda - u \in \mathcal{K}_\mathbb{R}, \\ D(W, \Theta) \in \mathbb{S}_{\succeq}^{2d}, (\lambda_i, v, \Theta_{i,i}) \in \mathcal{K}_{\log}, \forall i \in [\![d]\!]. \end{array}$$

**5.7. Polynomial weighted SOS and SOS matrix cones.** Our examples in subsections 6.6 to 6.8 use the EFs implicit in the definitions of $\mathcal{K}_{\text{sos}}^*$, $\mathcal{K}_{\text{sos}}$, and $\mathcal{K}_{\text{sosm}}$ in (4.23) and (4.25a).

**6. Numerical examples.** In subsections 6.1 to 6.8, we present example problems with NFs using some of Hypatia's predefined cones from section 4 and EFs constructed using the techniques from section 5. For each example problem, we generate random instances of a wide variety of sizes, and we observe larger dimensions and often larger barrier parameters for EFs compared to NFs. In Tables 6.1 to 6.11, $\nu$ and $n$, $p$, $q$ refer to the NF barrier parameter and primal variable, linear equality, and cone inequality dimensions (in our general conic form (3.2)), and $\bar{\nu}$, $\bar{n}$, $\bar{p}$, $\bar{q}$ refer to the corresponding EF values. For three solver/formulation combinations - Hypatia with NF (*Hypatia-NF*), Hypatia with EF (*Hypatia-EF*), and MOSEK with EF (*MOSEK-EF*) - we compare termination statuses, iteration counts, and solve times in seconds (columns *st*, *it*, and *time*) in Tables 6.1 to 6.11 and Figure 6.1. In subsections 6.2 and 6.5 we depend on a geometric mean cone EF, so we compare the *EF-exp* and *EF-sec* formulations from subsection 5.2. Note that all of our instances are primal-dual feasible, so we expect solvers to return optimality certificates. Compared to Hypatia-EF and MOSEK-EF, Hypatia-NF generally converges faster and more reliably, and solves larger instances within time and memory limits.

We use JuMP 0.21.5 and MathOptInterface 0.9.18 to build all models and interface with Hypatia 0.3.0 and MosekTools 0.9.4 and MOSEK 9. We perform all experiments on dedicated hardware with an AMD Ryzen 9 3950X 16-Core Processor (32 threads) and 128GB of RAM, running Ubuntu 20.10 and Julia 1.5.3, and we limit each solver to using 16 threads. MOSEK uses its conic interior point method for all problems, and we do not disable any MOSEK features. Hypatia and MOSEK use similar convergence criteria (see [23, section 13.3.2]), and we set their feasibility and optimality gap tolerances to $10^{-7}$. In the solver statistics tables, asterisks indicate missing data, and we use the following codes for the termination status (st) columns:

*co* - the solver claims it has an approximate optimality certificate,

*tl* - the solver stops itself due to a solve time limit of 1800 seconds, or the solve run is killed because it takes at least $1.2 \times 1800$ seconds,

*rl* - the solve is terminated because insufficient RAM is available,

*sp* - the solver reports *slow progress* during iterations,

*er* - the solver encounters a numerical error,

*m* - the model cannot be constructed with JuMP due to insufficient RAM or a model generation time limit of 3600 seconds (EFs tend to be slower and more memory-intensive to construct than NFs, so EF columns often have missing data),

*sk* - we skip the solve run because a smaller instance has a tl or rl status, or we skip model generation because a smaller instance has an m status.

For each solve run that yields a primal-dual point $(x, y, z, s)$ (see section 3; $s \in \mathcal{K}$ and $z \in \mathcal{K}^*$ are the solver's primal and dual cone interior points at termination), we compute:

$$(6.1) \qquad \epsilon = \max\left( \frac{\|A'y + G'z + c\|_\infty}{1 + \|c\|_\infty}, \frac{\|-Ax + b\|_\infty}{1 + \|b\|_\infty}, \frac{\|-Gx + h - s\|_\infty}{1 + \|h\|_\infty}, \frac{|c'x + b'y + h'z|}{1 + |b'y + h'z|} \right),$$

and if $\epsilon < 10^{-5}$, we underline the corresponding status code (e.g. $\underline{co}$, $\underline{tl}$) to indicate that the solution approximately satisfies the optimality certificate conditions from section 3. In our solve time plots in Figure 6.1, we only plot solve runs with underlined status codes. Finally, for each instance and each pair of corresponding solve runs with $\underline{co}$ status codes, we compute the relative difference of the primal objective values $g_1$ and $g_2$ as $\tilde{\epsilon} = |g_1 - g_2|/(1 + \max(|g_1|, |g_2|))$. We note $\tilde{\epsilon} < 10^{-4}$ in every case, and $\tilde{\epsilon} > 10^{-5}$ for only three instances in Table 6.11 (for $m \in \{10, 12, 14\}$).

**6.1. Portfolio rebalancing.** Suppose there are $k$ possible investments with expected returns $g \in \mathbb{R}_>^k$ and covariance matrix $\Sigma \in \mathbb{S}_\succeq^k$. We formulate a risk-constrained portfolio rebalancing optimization problem as:

$$(6.2a) \qquad \max_{\rho \in \mathbb{R}^k} \quad g'\rho :$$
$$(6.2b) \qquad e'\rho = 0,$$
$$(6.2c) \qquad (\gamma, \Sigma^{1/2}\rho) \in \mathcal{K}_{\ell_\infty},$$
$$(6.2d) \qquad (\gamma\sqrt{k}, \Sigma^{1/2}\rho) \in \mathcal{K}_{\ell_\infty}^*.$$

The EFs for (6.2c) and (6.2d) follow (5.1) and (5.2). Note the EF is a linear program.

To build random instances of (6.2), we generate $g$ with independent uniform positive entries and $\Sigma^{1/2}$ with independent Gaussian entries, for various values of $k$. We use $\Sigma$ to compute reasonable values for the risk parameter $\gamma > 0$. Our results are summarized in Table 6.1 and Figure 6.1a. The variable and conic constraint dimensions and barrier parameters of the EFs are approximately double those of the NFs. Hypatia-NF exhibits faster solve times and solves one instance larger than MOSEK-EF, which hits a RAM limit for $k = 11000$. Hypatia-EF encounters slow progress and a numerical error on moderate size instances. MOSEK requires notably fewer PDIPM iterations than Hypatia (which has no specializations for linear programs).

| | Hypatia-NF | | | Hypatia-EF | | | MOSEK-EF | | |
|---|---|---|---|---|---|---|---|---|---|
| $k$ | st | it | time | st | it | time | st | it | time |
| 500 | co | 38 | 0.4 | co | 43 | 1.1 | co | 18 | 0.7 |
| 1000 | co | 52 | 2.0 | co | 40 | 8.1 | co | 20 | 3.3 |
| 1500 | co | 56 | 5.6 | co | 57 | 30. | co | 23 | 9.7 |
| 2000 | co | 52 | 11. | co | 58 | 58. | co | 18 | 18. |
| 3000 | co | 74 | 43. | co | 60 | 168. | co | 25 | 58. |
| 4000 | co | 93 | 102. | sp | 15 | 161. | co | 29 | 121. |
| 5000 | co | 100 | 174. | er | 15 | 308. | co | 29 | 222. |
| 6000 | co | 122 | 324. | sp | 16 | 510. | co | 34 | 438. |
| 7000 | co | 134 | 506. | sp | 15 | 774. | co | 33 | 649. |
| 8000 | co | 127 | 694. | sp | 16 | 1156. | co | 30 | 767. |
| 9000 | co | 134 | 934. | tl | 22 | 1800. | co | 36 | 1330. |
| 10000 | co | 156 | 1399. | sk | * | * | co | 34 | 1505. |
| 11000 | co | 143 | 1625. | sk | * | * | rl | * | * |

Table 6.1: Portfolio rebalancing solver statistics. Note $\nu = q = 2k+2$, $\bar{\nu} = \bar{q} = 4k+1$, $n = k$, $\bar{n} = 2k$, $p = \bar{p} = 1$.

**6.2. Matrix completion.** Suppose there exists a matrix $F \in \mathbb{R}^{k \times l}$ and we know the entries $(F_{i,j})_{(i,j) \in \mathcal{S}}$ in the sparsity pattern $\mathcal{S}$. In the matrix completion problem, we seek to estimate the missing components $(F_{i,j})_{(i,j) \notin \mathcal{S}}$. We modify the formulation in [1, section 4.3] by replacing the spectral radius in the objective function with the spectral norm (allowing rectangular matrices) and using a convex relaxation of the geometric mean equality constraint:

$$(6.3a) \qquad \min_{\rho \in \mathbb{R}, X \in \mathbb{R}^{k \times l}} \quad \rho :$$
$$(6.3b) \qquad X_{i,j} = F_{i,j} \qquad\qquad \forall(i,j) \in \mathcal{S},$$
$$(6.3c) \qquad (\rho, \text{vec}(X)) \in \mathcal{K}_{\text{spec}(k,l)},$$
$$(6.3d) \qquad (1, (X_{i,j})_{(i,j) \notin \mathcal{S}}) \in \mathcal{K}_{\text{geo}}.$$

12

The EF for (6.3c) follows (5.5), and for (6.3d) we compare EF-exp and EF-sec (see subsection 5.2).

To build random instances of (6.3), we generate sparse matrices $F$ with independent Gaussian nonzero entries, for various values of $k$, column-to-row ratios $m \in \{10, 20\}$, and $l = mk$. Our results are summarized in Tables 6.2 and 6.3 and Figure 6.1b. Note we only plot EF-sec results for Hypatia-EF and MOSEK-EF, as MOSEK performs better with EF-sec than with EF-exp, though Hypatia exhibits the opposite trend. Hypatia-NF is much faster and solves more instances than the Hypatia-EFs and MOSEK-EFs, especially for the larger $m = 20$ value. The NFs have significantly lower barrier parameters than the EFs, and Hypatia-NF typically takes fewer than half the iterations compared to the Hypatia-EFs. Hypatia-NF usually also takes fewer iterations than MOSEK-EF-exp, but more than MOSEK-EF-sec (note EF-sec only uses symmetric cones, which may explain why MOSEK performs better with EF-sec than with EF-exp).

| $m$ | $k$ | NF | | | | EF-exp | | | EF-sec | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $\nu$ | $n$ | $p$ | $q$ | $\bar{\nu}$ | $\bar{n}$ | $\bar{q}$ | $\bar{\nu}$ | $\bar{n}$ | $\bar{q}$ |
| | 5 | 227 | 251 | 30 | 472 | 717 | 472 | 2202 | 566 | 506 | 2306 |
| | 10 | 900 | 1001 | 112 | 1890 | 2776 | 1890 | 8771 | 2157 | 2024 | 9175 |
| | 15 | 2032 | 2251 | 235 | 4267 | 6212 | 4267 | 19742 | 4260 | 4298 | 19837 |
| 10 | 20 | 3630 | 4001 | 392 | 7610 | 11046 | 7610 | 35136 | 8411 | 8096 | 36596 |
| | 25 | 5634 | 6251 | 643 | 11859 | 17098 | 11859 | 54773 | 16658 | 14442 | 62524 |
| | 30 | 8098 | 9001 | 934 | 17068 | 24530 | 17068 | 78815 | 16713 | 17192 | 79189 |
| | 35 | 11056 | 12251 | 1231 | 23271 | 33444 | 23271 | 107364 | 33152 | 28634 | 123455 |
| | 40 | 14444 | 16001 | 1598 | 30404 | 43648 | 30404 | 140228 | 33207 | 32384 | 146170 |
| | 5 | 451 | 501 | 56 | 946 | 1439 | 946 | 6899 | 1128 | 1012 | 7099 |
| | 10 | 1802 | 2001 | 210 | 3792 | 5582 | 3792 | 27527 | 4305 | 4048 | 28297 |
| 20 | 15 | 4069 | 4501 | 448 | 8554 | 12473 | 8554 | 61928 | 8506 | 8596 | 62056 |
| | 20 | 7200 | 8001 | 822 | 15180 | 21956 | 15180 | 109946 | 16803 | 16192 | 112984 |
| | 25 | 11270 | 12501 | 1257 | 23745 | 34256 | 23745 | 171806 | 33292 | 28884 | 187225 |
| | 30 | 16249 | 18001 | 1783 | 34219 | 49283 | 34219 | 247418 | 33397 | 34384 | 247915 |

Table 6.2: Matrix completion formulation statistics. Note $p = \bar{p} = |S|$.

**6.3. Multi-response regression.** In the multi-response linear regression problem, we seek to estimate a coefficient matrix $F \in \mathbb{R}^{m \times l}$ from a design matrix $X \in \mathbb{R}^{l \times k}$ and response matrix $Y \in \mathbb{R}^{m \times k}$. We use a similar formulation to the one proposed in [42], with nuclear norm loss and $\ell_2$ norm regularization:

$$(6.4a) \qquad \min_{\rho \in \mathbb{R}, \mu \in \mathbb{R}, F \in \mathbb{R}^{m \times l}} \quad \rho + \gamma\mu :$$
$$(6.4b) \qquad (\rho, \text{vec}(Y - FX)) \in \mathcal{K}^*_{\text{spec}(m,k)},$$
$$(6.4c) \qquad (\mu, \text{vec}(F)) \in \mathcal{K}_{\ell_2}.$$

The EF for NF constraint (6.4b) follows (5.6).

To build random instances of (6.4), we generate random $X$ and $Y$ with independent Gaussian entries, for various values of $k$ with $l = m \in \{15, 30\}$, and we use regularization parameter $\gamma = 0.1$. Our results are summarized in Table 6.4 and Figure 6.1c. Note that the variable dimensions and barrier parameters for the NFs only depend on $k$ and are much smaller than those of the EFs. The EFs also have much larger conic constraint dimensions. Hypatia-NF exhibits faster solve times than

| | | NF | | | EF-exp | | | | | | EF-sec | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Hypatia | | | Hypatia | | | MOSEK | | | Hypatia | | | MOSEK | | |
| $m$ | $k$ | st | it | time | st | it | time | st | it | time | st | it | time | st | it | time |
| | 5 | co | 14 | 0.1 | co | 27 | 0.9 | co | 15 | 0.9 | co | 28 | 1.0 | co | 9 | 0.6 |
| | 10 | co | 17 | 0.8 | co | 43 | 21. | co | 21 | 19. | co | 46 | 24. | co | 10 | 9.6 |
| | 15 | co | 21 | 8.6 | co | 49 | 118. | co | 20 | 100. | co | 47 | 128. | co | 12 | 65. |
| 10 | 20 | co | 21 | 29. | co | 56 | 487. | co | 24 | 472. | co | 73 | 803. | co | 13 | 266. |
| | 25 | co | 22 | 91. | co | 62 | 1736. | co | 25 | 1507. | tl | 40 | 1806. | co | 15 | 925. |
| | 30 | co | 23 | 253. | tl | 27 | 1808. | tl | 13 | 1903. | sk | * | * | tl | 12 | 1814. |
| | 35 | co | 26 | 619. | sk | * | * | sk | * | * | sk | * | * | sk | * | * |
| | 40 | co | 29 | 1407. | sk | * | * | sk | * | * | sk | * | * | sk | * | * |
| | 5 | co | 15 | 0.2 | co | 34 | 5.8 | co | 25 | 19. | co | 38 | 6.7 | co | 11 | 9.2 |
| | 10 | co | 19 | 5.9 | co | 49 | 139. | co | 26 | 414. | co | 60 | 207. | co | 15 | 249. |
| 20 | 15 | co | 22 | 40. | co | 59 | 1105. | tl | 16 | 1881. | co | 55 | 1063. | co | 14 | 1676. |
| | 20 | co | 23 | 186. | rl | * | * | sk | * | * | rl | * | * | rl | * | * |
| | 25 | co | 24 | 616. | sk | * | * | sk | * | * | sk | * | * | sk | * | * |
| | 30 | tl | 26 | 1803. | sk | * | * | sk | * | * | sk | * | * | sk | * | * |

Table 6.3: Matrix completion solver statistics.

Hypatia-EF and MOSEK-EF for almost all instances. Hypatia-NF solves much larger instances and takes a fairly consistent number of iterations.

| | | form. stats. | | Hypatia-NF | | | Hypatia-EF | | | MOSEK-EF | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $m$ | $k$ | $\bar{n}$ | $q$ | st | it | time | st | it | time | st | it | time |
| | 50 | 1622 | 977 | co | 11 | 0.2 | co | 11 | 1.3 | co | 4 | 0.6 |
| | 100 | 5397 | 1727 | co | 11 | 1.0 | co | 12 | 30. | co | 5 | 6.8 |
| | 150 | 11672 | 2477 | co | 11 | 2.0 | co | 13 | 240. | co | 5 | 36. |
| | 250 | 31722 | 3977 | co | 12 | 6.7 | tl | * | * | co | 5 | 319. |
| | 500 | 125597 | 7727 | co | 12 | 33. | sk | * | * | tl | * | * |
| | 750 | * | 11477 | co | 12 | 86. | sk | * | * | sk | * | * |
| 15 | 1000 | * | 15227 | co | 12 | 162. | sk | * | * | sk | * | * |
| | 1250 | * | 18977 | co | 14 | 304. | sk | * | * | sk | * | * |
| | 1500 | * | 22727 | co | 12 | 425. | sk | * | * | sk | * | * |
| | 1750 | * | 26477 | co | 12 | 620. | sk | * | * | sk | * | * |
| | 2000 | * | 30227 | co | 12 | 932. | sk | * | * | sk | * | * |
| | 2250 | * | 33977 | co | 12 | 1128. | sk | * | * | sk | * | * |
| | 50 | 2642 | 2402 | co | 11 | 1.5 | co | 10 | 4.3 | co | 4 | 1.2 |
| | 100 | 6417 | 3902 | co | 12 | 5.6 | co | 12 | 54. | co | 5 | 13. |
| | 150 | 12692 | 5402 | co | 13 | 13. | co | 13 | 351. | co | 5 | 50. |
| 30 | 250 | 32742 | 8402 | co | 13 | 42. | tl | * | * | co | 5 | 427. |
| | 500 | 126617 | 15902 | co | 13 | 214. | sk | * | * | rl | * | * |
| | 750 | * | 23402 | co | 13 | 525. | sk | * | * | sk | * | * |
| | 1000 | * | 30902 | co | 13 | 1155. | sk | * | * | sk | * | * |
| | 1250 | * | 38402 | co | 13 | 1851. | sk | * | * | sk | * | * |

Table 6.4: Multi-response regression formulation and solver statistics. Note $\nu = 3+m$, $\bar{\nu} = \nu + k$, $n = 2 + m^2$, $p = \bar{p} = 0$, $\bar{q} = \bar{n} + mk$.

**6.4. Sparse/completable PSD matrix.** Given a symmetric indefinite matrix $H \in \mathbb{S}^k$ with sparsity pattern $\mathcal{S}$, we consider two problem variants that seek a symmetric matrix $F$ with sparsity pattern $\mathcal{S}$ that maximizes the inner product with $H$

subject to a normalization constraint. In the sparse PSD variant (*sparse*), $F$ must be PSD, and in the PSD completion variant (*compl*), $F$ must have a PSD completion (see [4]). Using the operator $\mathrm{mat}_\mathcal{S}$ defined in (4.16), we formulate these problems as:

$$(6.5a) \qquad \max_{f \in \mathbb{R}^{|\mathcal{S}|}} \quad \mathrm{tr}(H \, \mathrm{mat}_\mathcal{S}(f)) :$$

$$(6.5b) \qquad \qquad \mathrm{tr}(\mathrm{mat}_\mathcal{S}(f)) = 1,$$

$$(6.5c) \qquad \qquad f \in \mathcal{K},$$

where $\mathcal{K}$ in (6.5c) is $\mathcal{K}_{\mathrm{spsd}(\mathcal{S})}$ for the sparse variant or $\mathcal{K}^*_{\mathrm{spsd}(\mathcal{S})}$ for the compl variant. Subsection 5.5 discusses the EFs for each variant of (6.5c).

To build random instances of (6.5), we generate sparse matrices $H$ with nonzero diagonal and independent Gaussian entries, for various side dimensions $k$. Our results are summarized in Table 6.5 and Figure 6.1d. The NFs and EFs have the same barrier parameters, and Hypatia-NF and Hypatia-EF take similar numbers of iterations. For both variants, the EFs have much higher cone dimensions, and for the compl variant, the EFs also have much higher variable dimensions. Hypatia-NF is faster than MOSEK-EF except for the smallest instances, and Hypatia-NF solves larger instances than Hypatia-EF and MOSEK-EF. MOSEK-EF and Hypatia-NF each perform similarly for the sparse variant compared to the compl variant. Compared to Hypatia-NF and MOSEK-EF, Hypatia-EF is much faster on the sparse variant (though it hits a RAM limit at $k = 500$) but much slower on the compl variant (the divergent behavior of MOSEK-EF and Hypatia-EF may be attributed to different algorithmic choices).

| | | form. stats. | | Hypatia-NF | | | Hypatia-EF | | | MOSEK-EF | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| variant | $k$ | $n$ | $\bar{q}$ | st | it | time | st | it | time | st | it | time |
| | 50 | 126 | 1275 | co | 11 | 1.2 | co | 11 | 0.1 | co | 5 | 0.2 |
| | 100 | 261 | 5050 | co | 14 | 9.0 | co | 14 | 0.5 | co | 5 | 4.2 |
| | 150 | 370 | 11325 | co | 15 | 18. | co | 15 | 1.6 | co | 5 | 25. |
| | 200 | 511 | 20100 | co | 15 | 51. | co | 16 | 5.0 | co | 5 | 94. |
| sparse | 300 | 747 | 45150 | co | 17 | 143. | co | 17 | 22. | co | 5 | 606. |
| | 400 | 1011 | 80200 | co | 17 | 339. | co | 17 | 57. | rl | * | * |
| | 500 | 1248 | 125250 | co | 18 | 590. | rl | * | * | sk | * | * |
| | 600 | 1506 | 180300 | co | 18 | 996. | sk | * | * | sk | * | * |
| | 700 | 1761 | 245350 | tl | 18 | 1843. | sk | * | * | sk | * | * |
| | 50 | 126 | 1275 | co | 10 | 1.0 | co | 9 | 0.9 | co | 5 | 0.2 |
| | 100 | 261 | 5050 | co | 13 | 7.9 | co | 13 | 20. | co | 5 | 4.5 |
| | 150 | 370 | 11325 | co | 13 | 13. | co | 14 | 194. | co | 5 | 24. |
| | 200 | 511 | 20100 | co | 15 | 41. | co | 14 | 968. | co | 5 | 90. |
| compl | 300 | 747 | 45150 | co | 15 | 103. | tl | * | * | co | 6 | 709. |
| | 400 | 1011 | 80200 | co | 15 | 241. | sk | * | * | rl | * | * |
| | 500 | 1248 | 125250 | co | 14 | 391. | m | * | * | sk | * | * |
| | 600 | 1506 | 180300 | co | 16 | 725. | sk | * | * | sk | * | * |
| | 700 | 1761 | 245350 | co | 17 | 1296. | sk | * | * | sk | * | * |

Table 6.5: Sparse/completable PSD matrix formulation and solver statistics. Note $\nu = \bar{\nu} = k$, $n = q$, $p = \bar{p} = 1$, and for the sparse variant, $\bar{n} = n$, and for the compl variant, $\bar{n} = \bar{q}$.

**6.5. D-optimal experiment design.** In a continuous relaxation of the D-optimal experiment design problem (see [9, section 7.5]), the variable $\mu \in \mathbb{R}^m$ is

the number of trials to run for each of $m$ experiments, and our goal is to minimize the determinant of the error covariance matrix $(F \operatorname{Diag}(\mu)F')^{-1}$, given a menu of experiments $F \in \mathbb{R}^{k \times m}$ useful for estimating a vector in $\mathbb{R}^k$. We require that a total of $j$ experiments are performed and that each experiment can be performed between 0 and $l$ times. We formulate this problem as:

$$\text{(6.6a)} \qquad \max_{\rho \in \mathbb{R}, \mu \in \mathbb{R}^m} \quad \rho :$$
$$\text{(6.6b)} \qquad e'\mu = j,$$
$$\text{(6.6c)} \qquad (\tfrac{l}{2}, \mu - \tfrac{l}{2}e) \in \mathcal{K}_{\ell_\infty},$$
$$\text{(6.6d)} \qquad (\rho, \operatorname{vec}(F \operatorname{Diag}(\mu)F')) \in \mathcal{K}_{\text{rtdet}}.$$

In an alternative *logdet* variant of the *rtdet* variant (6.6), we replace (6.6d) with:

$$\text{(6.7)} \qquad (\rho, 1, \operatorname{vec}(F \operatorname{Diag}(\mu)F')) \in \mathcal{K}_{\text{logdet}},$$

noting that both variants have the same optimal solution set for $\mu$. The EFs for (6.6c), (6.6d), and (6.7) follow (5.1), (5.7), and (5.8). Since the EF for $\mathcal{K}_{\text{rtdet}}$ depends on a $\mathcal{K}_{\text{geo}}$ EF, for the rtdet variant we compare EF-exp and EF-sec (see subsection 5.2).

To build random instances of (6.6), we generate $F$ with independent Gaussian entries, for various values of $k$, $m = j = 2k$, and $l = 5$. Our results are summarized in Tables 6.6 to 6.8 and Figure 6.1e. Note for the rootdet variant, we only plot EF-sec results for Hypatia-EF and MOSEK-EF, as MOSEK typically performs slightly better with EF-sec than with EF-exp, though Hypatia exhibits the opposite trend. For both variants, the NFs have much lower barrier parameters and variable and conic constraint dimensions than the EFs. The Hypatia-NFs take fewer iterations than the Hypatia-EFs, particularly for rootdet. Hypatia-EF and MOSEK-EF have similar solve times, though MOSEK-EF takes fewer iterations. Although the EF solvers typically solve instances up to $k = 150$, Hypatia-NF solves instances with $k$ around twice that. Hypatia-NF is also much faster than the EF solvers for all $k$.

| | form. stats. | | | Hypatia-NF | | | Hypatia-EF | | | MOSEK-EF | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $k$ | $\bar{n}$ | $q$ | $\bar{q}$ | st | it | time | st | it | time | st | it | time |
| 25 | 401 | 378 | 1451 | co | 21 | 0.1 | co | 22 | 0.5 | co | 16 | 0.6 |
| 50 | 1426 | 1378 | 5401 | co | 25 | 0.4 | co | 25 | 5.0 | co | 14 | 11. |
| 75 | 3076 | 3003 | 11851 | co | 29 | 1.9 | co | 31 | 39. | co | 16 | 68. |
| 100 | 5351 | 5253 | 20801 | co | 29 | 7.4 | co | 31 | 157. | co | 15 | 255. |
| 125 | 8251 | 8128 | 32251 | co | 32 | 19. | co | 34 | 531. | co | 15 | 676. |
| 150 | 11776 | 11628 | 46201 | co | 32 | 54. | co | 36 | 1438. | co | 14 | 1668. |
| 175 | 15926 | 15753 | 62651 | co | 33 | 106. | tl | 4 | 1847. | tl | 6 | 1856. |
| 200 | 20701 | 20503 | 81601 | co | 34 | 213. | sk | * | * | sk | * | * |
| 225 | 26101 | 25878 | 103051 | co | 34 | 429. | sk | * | * | sk | * | * |
| 250 | 32126 | 31878 | 127001 | co | 35 | 720. | sk | * | * | sk | * | * |
| 275 | 38776 | 38503 | 153451 | co | 37 | 1265. | sk | * | * | sk | * | * |
| 300 | 46051 | 45753 | 182401 | tl | 33 | 1832. | sk | * | * | sk | * | * |

Table 6.6: D-optimal experiment design logdet variant formulation and solver statistics. Note $\nu = 3 + 3k$, $\bar{\nu} = 1 + 9k$, $n = 1 + 2k$, $p = \bar{p} = 1$.

**6.6. Polynomial minimization.** Following [33], we use an interpolant basis weighted SOS dual formulation to find a lower bound for a multivariate polynomial

|  | NF | | | EF-exp | | | EF-sec | | |
|---|---|---|---|---|---|---|---|---|---|
| $k$ | $\nu$ | $n$ | $q$ | $\bar{\nu}$ | $\bar{n}$ | $\bar{q}$ | $\bar{\nu}$ | $\bar{n}$ | $\bar{q}$ |
| 25 | 123 | 51 | 377 | 227 | 402 | 1452 | 213 | 407 | 1469 |
| 50 | 243 | 101 | 1377 | 452 | 1427 | 5402 | 427 | 1439 | 5440 |
| 75 | 362 | 151 | 3002 | 677 | 3077 | 11852 | 705 | 3128 | 12007 |
| 100 | 482 | 201 | 5252 | 902 | 5352 | 20802 | 855 | 5378 | 20882 |
| 125 | 601 | 251 | 8127 | 1127 | 8253 | 32252 | 1005 | 8253 | 32257 |
| 150 | 720 | 301 | 11627 | 1352 | 11777 | 46202 | 1411 | 11881 | 46516 |
| 175 | 840 | 351 | 15752 | 1577 | 15927 | 62652 | 1561 | 16006 | 62891 |
| 200 | 959 | 401 | 20502 | 1802 | 20702 | 81602 | 1711 | 20756 | 81766 |
| 225 | 1079 | 451 | 25877 | 2027 | 26102 | 103052 | 1861 | 26131 | 103141 |
| 250 | 1198 | 501 | 31877 | 2252 | 32127 | 127002 | 2011 | 32131 | 127016 |
| 275 | 1318 | 551 | 38502 | 2477 | 38777 | 153452 | 2673 | 39012 | 154159 |
| 300 | 1437 | 601 | 45752 | 2702 | 46052 | 182402 | 2823 | 46262 | 183034 |

Table 6.7: D-optimal experiment design rtdet variant formulation statistics. Note $p = \bar{p} = 1$ and $\nu$ is rounded to the nearest integer.

|  | NF | | | EF-exp | | | | | | EF-sec | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
|  | Hypatia | | | Hypatia | | | MOSEK | | | Hypatia | | | MOSEK | | |
| $k$ | st | it | time | st | it | time | st | it | time | st | it | time | st | it | time |
| 25 | co | 17 | 0.0 | co | 22 | 0.5 | co | 15 | 0.6 | co | 25 | 0.6 | co | 12 | 0.6 |
| 50 | co | 20 | 0.3 | co | 25 | 5.1 | co | 12 | 9.2 | co | 28 | 5.8 | co | 11 | 9.7 |
| 75 | co | 24 | 1.6 | co | 30 | 38. | co | 15 | 64. | co | 35 | 45. | co | 13 | 60. |
| 100 | co | 24 | 6.3 | co | 31 | 156. | co | 13 | 229. | co | 35 | 172. | co | 10 | 194. |
| 125 | co | 25 | 15. | co | 33 | 514. | co | 13 | 603. | co | 37 | 552. | co | 11 | 553. |
| 150 | co | 26 | 39. | co | 35 | 1419. | co | 12 | 1462. | sp | 45 | 1654. | co | 11 | 1396. |
| 175 | co | 27 | 84. | tl | 4 | 1848. | tl | 6 | 1885. | tl | 3 | 1826. | tl | 6 | 1904. |
| 200 | co | 29 | 197. | sk | * | * | sk | * | * | sk | * | * | sk | * | * |
| 225 | co | 29 | 367. | sk | * | * | sk | * | * | sk | * | * | sk | * | * |
| 250 | co | 29 | 591. | sk | * | * | sk | * | * | sk | * | * | sk | * | * |
| 275 | co | 30 | 1131. | sk | * | * | sk | * | * | sk | * | * | sk | * | * |
| 300 | tl | 32 | 1805. | sk | * | * | sk | * | * | sk | * | * | sk | * | * |

Table 6.8: D-optimal experiment design rtdet variant solver statistics.

$f$ of maximum degree $2k$ in $m$ variables over the unit hypercube $\mathcal{D} = [-1, 1]^m$. We let $U = \binom{m+2k}{m}$, $L = \binom{m+k}{m}$, $\tilde{L} = \binom{m+k-1}{m}$. We select multivariate Chebyshev basis polynomials $g_j, \forall j \in \llbracket L \rrbracket$ of increasing degree up to $k$, and suitable interpolation points $o_u \in \mathcal{D}, \forall u \in \llbracket U \rrbracket$. To parametrize $\mathcal{K}^*_{\text{sos}(P)}$, we set up the collection of matrices $P$ by evaluating functions of basis polynomials at the points:

(6.8a) $\qquad (P_1)_{u,j} = g_j(o_u) \qquad\qquad\qquad\qquad \forall u \in \llbracket U \rrbracket, j \in \llbracket L \rrbracket,$

(6.8b) $\qquad (P_{1+i})_{u,j} = g_j(o_u)(1 - o_{u,i}^2) \qquad \forall i \in \llbracket m \rrbracket, u \in \llbracket U \rrbracket, j \in \llbracket \tilde{L} \rrbracket.$

Letting $\bar{f} = (f(o_u))_{u \in U}$ be evaluations of $f$ at the points, the conic formulation is:

(6.9a) $\qquad\qquad\qquad \min_{\rho \in \mathbb{R}^U} \quad \bar{f}'\rho :$

(6.9b) $\qquad\qquad\qquad\qquad\qquad e'\rho = 1,$

(6.9c) $\qquad\qquad\qquad\qquad\qquad \rho \in \mathcal{K}^*_{\text{sos}(P)}.$

The EF for NF constraint (6.9c) follows subsection 5.7.

To build random instances of (6.9), we generate $\bar{f}$ (which implicitly defines a polynomial $f$) with independent Gaussian entries, for various values of $m$ and $k$. Our results are summarized in Table 6.9 and Figure 6.1f. In order to show interesting trend lines, we only plot results for Hypatia-NF on instances with $m \leq 4$. The NFs and EFs have the same barrier parameters and variable and equality dimensions. For fixed $m$, the conic constraint dimensions are larger for the EFs and grow much faster for the EFs as the degree $k$ increases. Hypatia-NF is faster than the EF solvers on instances with $k > 2$, and solves instances with much higher degrees.

| | | form. stats. | | | Hypatia-NF | | | Hypatia-EF | | | MOSEK-EF | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $m$ | $k$ | $\nu$ | $n$ | $\bar{q}$ | st | it | time | st | it | time | st | it | time |
| 1 | 100 | 201 | 201 | 10201 | co | 16 | 0.1 | co | 32 | 1.7 | co | 14 | 25. |
| 1 | 200 | 401 | 401 | 40401 | co | 17 | 0.4 | co | 44 | 22. | co | 15 | 588. |
| 1 | 500 | 1001 | 1001 | 251001 | co | 20 | 3.6 | rl | * | * | rl | * | * |
| 1 | 1000 | 2001 | 2001 | * | co | 23 | 19. | m | * | * | m | * | * |
| 1 | 2000 | 4001 | 4001 | * | co | 27 | 135. | sk | * | * | sk | * | * |
| 1 | 3000 | 6001 | 6001 | * | co | 29 | 440. | sk | * | * | sk | * | * |
| 1 | 4000 | 8001 | 8001 | * | co | 30 | 978. | sk | * | * | sk | * | * |
| 2 | 15 | 376 | 496 | 23836 | co | 21 | 0.6 | co | 24 | 7.5 | co | 9 | 99. |
| 2 | 30 | 1426 | 1891 | 339946 | co | 27 | 17. | rl | * | * | rl | * | * |
| 2 | 45 | 3151 | 4186 | * | co | 30 | 136. | m | * | * | m | * | * |
| 2 | 60 | 5551 | 7381 | * | co | 33 | 651. | sk | * | * | sk | * | * |
| 3 | 6 | 252 | 455 | 8358 | co | 22 | 0.5 | co | 18 | 2.0 | co | 8 | 7.7 |
| 3 | 9 | 715 | 1330 | 65395 | co | 25 | 6.0 | co | 25 | 73. | co | 8 | 675. |
| 3 | 12 | 1547 | 2925 | 303030 | co | 29 | 40. | rl | * | * | rl | * | * |
| 3 | 15 | 2856 | 5456 | * | co | 32 | 211. | m | * | * | m | * | * |
| 4 | 4 | 210 | 495 | 5005 | co | 21 | 1.3 | co | 17 | 2.2 | co | 7 | 3.1 |
| 4 | 6 | 714 | 1820 | 54159 | co | 23 | 9.6 | co | 19 | 72. | co | 14 | 624. |
| 4 | 8 | 1815 | 4845 | * | co | 23 | 100. | m | * | * | m | * | * |
| 4 | 10 | 3861 | 10626 | * | co | 29 | 867. | sk | * | * | sk | * | * |
| 8 | 2 | 117 | 495 | 1395 | co | 21 | 1.3 | co | 15 | 0.7 | co | 6 | 0.5 |
| 8 | 3 | 525 | 3003 | 21975 | co | 27 | 34. | co | 19 | 71. | co | 11 | 162. |
| 8 | 4 | 1815 | 12870 | * | co | 34 | 1505. | m | * | * | m | * | * |
| 16 | 1 | 33 | 153 | 169 | co | 17 | 0.1 | co | 15 | 0.7 | co | 8 | 0.0 |
| 16 | 2 | 425 | 4845 | 14229 | co | 32 | 139. | co | 24 | 96. | co | 9 | 160. |
| 32 | 1 | 65 | 561 | 593 | co | 17 | 1.1 | co | 13 | 1.2 | co | 8 | 0.2 |
| 64 | 1 | 129 | 2145 | 2209 | co | 20 | 31. | co | 15 | 3.4 | co | 9 | 2.9 |

Table 6.9: Polynomial minimization formulation and solver statistics. Note $\nu = \bar{\nu}$, $p = \bar{p} = 1$, $n = \bar{n} = q$.

**6.7. Smooth density estimation.** $\mathbb{R}_{m,2k}[x]$ is the ring of polynomials of maximum degree $2k$ in $m$ variables [33]. We seek a polynomial density function $f \in \mathbb{R}_{m,2k}[x]$ over the domain $\mathcal{D} = [-1, 1]^m$ that maximizes the log-likelihood of $N$ given observations $z_i \in \mathcal{D}, \forall i \in [\![N]\!]$ (compare to [31, section 4.3]). For $f$ to be a valid density it must be nonnegative on $\mathcal{D}$ and integrate to one over $\mathcal{D}$, so we aim to solve:

$$(6.10a) \qquad \max_{f \in \mathbb{R}_{m,2k}[x]} \quad \sum_{i \in [\![N]\!]} \log(f(z_i)) :$$

$$(6.10b) \qquad \qquad \int_{\mathcal{D}} f(x)\, dx = 1,$$

$$(6.10c) \qquad \qquad f(x) \geq 0 \qquad \qquad \forall x \in \mathcal{D}.$$

To find a feasible solution for (6.10), we build an SOS formulation. We obtain interpolation points and matrices $P$ parametrizing $\mathcal{K}_{\text{sos}(P)}$, using the techniques from subsection 6.6. From the interpolation points and the domain $\mathcal{D}$, we compute a vector of quadrature weights $\mu \in \mathbb{R}^U$. We compute a matrix $B \in \mathbb{R}^{N \times U}$ by evaluating the $U$ Lagrange basis polynomials corresponding to the interpolation points (see [33]) at the $N$ observations. Letting variable $\rho$ represent the coefficients on the Lagrange basis, the conic formulation is:

$$
\text{(6.11a)} \qquad \max_{\psi \in \mathbb{R}, \rho \in \mathbb{R}^U} \quad \psi :
$$

$$
\text{(6.11b)} \qquad \qquad \qquad \mu' \rho = 1,
$$

$$
\text{(6.11c)} \qquad \qquad \qquad (\psi, 1, B\rho) \in \mathcal{K}_{\log},
$$

$$
\text{(6.11d)} \qquad \qquad \qquad \rho \in \mathcal{K}_{\text{sos}(P)}.
$$

The EFs for NF constraints (6.11c) and (6.11d) follow subsections 5.3 and 5.7.

To build random instances of (6.11) for various values of $m$ and $k$, we generate $N = 500$ independent uniform samples in $[-1,1]^m$ for $z_i \in \mathcal{D}, \forall i \in [\![N]\!]$. As our method for computing $\mu$ is numerically unstable for larger $m$, we only use $m \leq 16$. Our results are summarized in Table 6.10. The barrier parameters and all dimensions are larger for the EFs than for the NFs, and the EFs often hit RAM limits during model generation. The instances are numerically challenging, and MOSEK-EF often encounters slow progress. For reasons unknown, Hypatia-NF usually takes more iterations than Hypatia-EF on the instances solved by both. Hypatia-NF is faster than the EF solvers on instances with $2k > 4$, and solves instances with much higher degrees.

**6.8. Shape constrained regression.** A common type of shape constraint imposes monotonicity or convexity of a polynomial over a basic semialgebraic set [18, section 6]. Given an $m$-dimensional feature variable $z$ and a scalar response variable $g$, we aim to fit a polynomial $f \in \mathbb{R}_{m,2k}[x]$ that is convex over $\mathcal{D} = [-1,1]^m$ to $N$ given observations $(z_i, g_i)_{i \in [\![N]\!]}$ with $z_i \in \mathcal{D}, \forall i \in [\![N]\!]$:

$$
\text{(6.12a)} \qquad \min_{f \in \mathbb{R}_{m,2k}[x]} \quad \sum_{i \in [\![N]\!]} (g_i - f(z_i))^2 :
$$

$$
\text{(6.12b)} \qquad \qquad y'(\nabla^2 f(x))y \geq 0 \qquad \forall x \in \mathcal{D}, y \in \mathbb{R}^m.
$$

Constraint (6.12b) ensures the Hessian matrix $\nabla^2 f(x)$ is PSD at every point $x \in \mathcal{D}$, which is equivalent to convexity of $f$ over $\mathcal{D}$. To find a feasible solution for (6.12), we build an SOS formulation. The polynomial variable, represented in an interpolant basis with the optimization variable $\rho \in \mathbb{R}^U$, has degree $2k$ and $U = \binom{m+2k}{m}$ coefficients. Each polynomial entry of $\nabla^2 f(x)$ has degree $2k - 2$ and $\bar{U} = \binom{m+2k-2}{m}$ coefficients. Following the descriptions in subsections 6.6 to 6.7, we obtain interpolation points and a Lagrange polynomial basis for these $U$-dimensional and $\bar{U}$-dimensional spaces, and we define the matrix $B \in \mathbb{R}^{N \times U}$ containing evaluations of the $U$-dimensional Lagrange basis at the $N$ feature observations. Finally, we let $F \in \mathbb{R}^{\text{sd}(m)\bar{U} \times U}$ be such that $F\rho$ is a vectorization of the tensor $H \in \mathbb{R}^{m \times m \times \bar{U}}$ (scaled to account for symmetry) with $H_{a,b,u}$ equal to the $u$th coefficient of the $(a,b)$th polynomial in $\nabla^2 f(x)$ for $a, b \in [\![m]\!]$ and $u \in [\![\bar{U}]\!]$. This yields the formulation:

$$
\text{(6.13a)} \qquad \min_{\psi \in \mathbb{R}, \rho \in \mathbb{R}^U} \quad \psi :
$$

$$
\text{(6.13b)} \qquad \qquad (\psi, g - B\rho) \in \mathcal{K}_{\ell_2},
$$

$$
\text{(6.13c)} \qquad \qquad F\rho \in \mathcal{K}_{\text{sosm}(P)}.
$$

| | | dimensions | | | Hypatia-NF | | | Hypatia-EF | | | MOSEK-EF | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $m$ | $2k$ | $\nu$ | $n$ | $\bar{n}$ | st | it | time | st | it | time | st | it | time |
| 1 | 100 | 603 | 102 | 3203 | co | 42 | 0.8 | co | 29 | 19. | co | 18 | 2.2 |
| 1 | 200 | 703 | 202 | 10903 | co | 45 | 0.9 | sp | 52 | 744. | co | 26 | 43. |
| 1 | 500 | 1003 | 502 | 64003 | co | 42 | 1.8 | rl | * | * | tl | 21 | 1859. |
| 1 | 1000 | 1503 | 1002 | * | co | 50 | 6.3 | sk | * | * | sk | * | * |
| 1 | 2000 | 2503 | 2002 | * | co | 56 | 29. | sk | * | * | sk | * | * |
| 1 | 3000 | 3503 | 3002 | * | co | 78 | 110. | sk | * | * | sk | * | * |
| 1 | 4000 | 4503 | 4002 | * | co | 96 | 261. | sk | * | * | sk | * | * |
| 1 | 5000 | 5503 | 5002 | * | co | 117 | 582. | sk | * | * | sk | * | * |
| 2 | 20 | 678 | 232 | 6023 | co | 57 | 0.4 | co | 35 | 83. | co | 22 | 8.0 |
| 2 | 40 | 1153 | 862 | 72468 | co | 51 | 4.3 | rl | * | * | sp | 23 | 1396. |
| 2 | 60 | 1928 | 1892 | * | co | 49 | 21. | sk | * | * | m | * | * |
| 2 | 80 | 3003 | 3322 | * | co | 62 | 96. | sk | * | * | sk | * | * |
| 2 | 100 | 4378 | 5152 | * | co | 76 | 376. | sk | * | * | sk | * | * |
| 3 | 12 | 754 | 456 | 9314 | co | 66 | 1.4 | co | 33 | 207. | sp | 22 | 21. |
| 3 | 18 | 1217 | 1331 | * | co | 67 | 12. | m | * | * | m | * | * |
| 3 | 24 | 2049 | 2926 | * | co | 68 | 81. | sk | * | * | sk | * | * |
| 3 | 30 | 3358 | 5457 | * | co | 79 | 473. | sk | * | * | sk | * | * |
| 4 | 8 | 712 | 496 | 6001 | co | 81 | 2.0 | co | 36 | 95. | sp | 20 | 9.4 |
| 4 | 12 | 1216 | 1821 | * | co | 93 | 34. | tl | * | * | m | * | * |
| 4 | 16 | 2317 | 4846 | * | co | 116 | 488. | sk | * | * | sk | * | * |
| 8 | 4 | 619 | 496 | 2391 | co | 85 | 1.9 | co | 40 | 11. | sp | 16 | 1.9 |
| 8 | 6 | 1027 | 3004 | * | co | 115 | 126. | tl | 21 | 1820. | m | * | * |
| 16 | 2 | 535 | 154 | 823 | co | 87 | 0.6 | co | 52 | 2.0 | sp | 24 | 0.2 |
| 16 | 4 | 927 | 4846 | * | sp | 128 | 697. | co | 47 | 1004. | m | * | * |

Table 6.10: Smooth density estimation. Note $\bar{\nu} = 999 + \nu$, $p = 1$, $\bar{p} = n$, $q = 501 + n$, $\bar{q} = 1001 + \bar{n} - n$.

Note that for $N > U$, we use a QR factorization to reduce the dimension of $\mathcal{K}_{\ell_2}$ in (6.13b) from $1+N$ to $2+U$.[4] The EF for NF constraint (6.13c) follows subsection 5.7.

To build random instances of (6.13) for various values of $m$ and $k$, we generate $N = \lceil 1.1U \rceil$ independent observations with $z_i$ sampled uniformly from $\mathcal{D}$ and $g_i = \exp(\|z\|^2/m) - 1 + \varepsilon_i$, where $\varepsilon_i$ is a Gaussian sample yielding a signal to noise ratio of 10, for all $i \in [\![N]\!]$. Note we exclude the case $m = 1$, since $\mathcal{K}_{\mathrm{sos}(P)}$ could be used in place of $\mathcal{K}_{\mathrm{sosm}(P)}$. Our results are summarized in Table 6.11. The barrier parameters are the same for EFs and NFs, but all dimensions are larger for the EFs, and the EFs often hit RAM limits during model generation. The instances are numerically challenging, and MOSEK-EF often encounters slow progress. Hypatia-NF usually takes fewer iterations than Hypatia-EF. Hypatia-NF is faster than the EF solvers on instances with $2k > 4$, and solves instances with much higher degrees.

**7. Conclusions.** Although many convex problems are representable with conic EFs using the small number of standard cones currently recognized by some advanced conic solvers, these formulations can be much larger than NFs with exotic cones. Using Hypatia's generic cone interface, we implement a variety of exotic cones, some of which are novel. We describe general techniques for constructing EFs and analyze how these techniques can increase the barrier parameter and variable, equality, and conic constraint dimensions associated with a formulation. For several example problems,

---

[4]Let $[-B \ \ g] = QR$, where $Q \in \mathbb{R}^{N \times (U+1)}$ has orthonormal columns and $R \in \mathbb{R}^{(U+1) \times (U+1)}$ is upper triangular. Then $(\psi, g - B\rho) \in \mathcal{K}_{\ell_2}$ if and only if $(\psi, R(\rho, 1)) \in \mathcal{K}_{\ell_2}$.
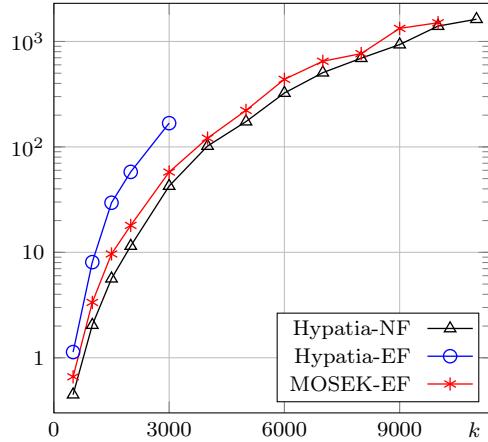
(a) 6.1: Portfolio rebalancing.
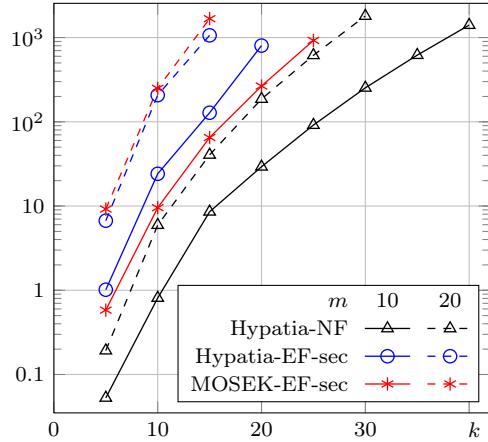
(b) 6.2: Matrix completion.
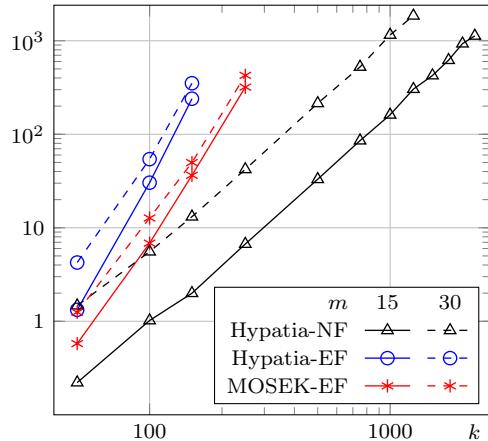
(c) 6.3: Multi-response regression (log-log).
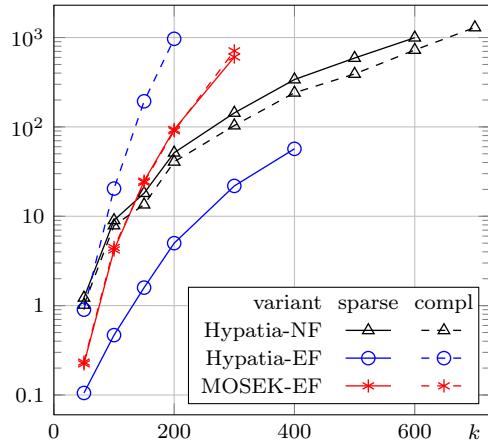
(d) 6.4: Sparse/completable PSD matrix.
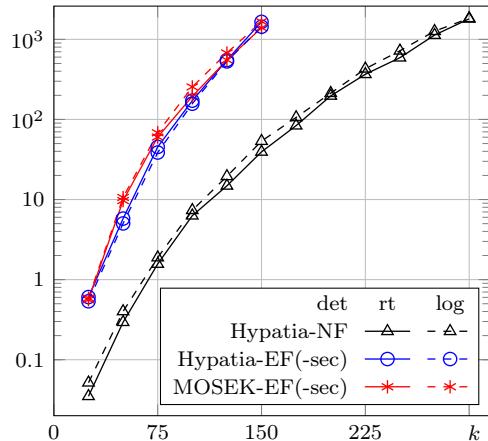
(e) 6.5: D-optimal experiment design.
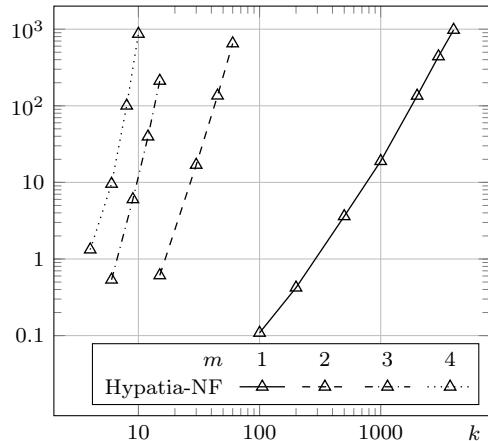
(f) 6.6: Polynomial minimization (log-log).

Fig. 6.1: Solve times (in seconds) for solve runs satisfying the convergence check (6.1).

| | | form. stats. | | | | Hypatia-NF | | | Hypatia-EF | | | MOSEK-EF | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $m$ | $2k$ | $\nu$ | $n$ | $\bar{n}$ | $q$ | st | it | time | st | it | time | st | it | time |
| 2 | 10 | 72 | 67 | 952 | 203 | co | 23 | 0.1 | co | 31 | 1.1 | sp | 14 | 0.3 |
| 2 | 20 | 292 | 232 | 14527 | 803 | co | 45 | 1.2 | co | 61 | 1033. | sp | 18 | 56. |
| 2 | 30 | 662 | 497 | 73727 | 1803 | co | 70 | 15. | rl | * | * | tl | 13 | 1902. |
| 2 | 40 | 1182 | 862 | 234052 | 3203 | co | 96 | 89. | m | * | * | sk | * | * |
| 2 | 50 | 1852 | 1327 | 573502 | 5003 | co | 115 | 363. | sk | * | * | sk | * | * |
| 2 | 60 | 2672 | 1892 | * | 7203 | co | 145 | 1412. | sk | * | * | m | * | * |
| 3 | 8 | 152 | 166 | 3391 | 671 | co | 25 | 0.4 | co | 36 | 15. | sp | 18 | 3.3 |
| 3 | 12 | 485 | 456 | 31347 | 2173 | co | 46 | 7.4 | tl | * | * | sp | 19 | 273. |
| 3 | 16 | 1118 | 970 | 161584 | 5051 | co | 69 | 84. | m | * | * | rl | * | * |
| 3 | 20 | 2147 | 1772 | 588182 | 9753 | co | 95 | 579. | sk | * | * | sk | * | * |
| 4 | 6 | 142 | 211 | 2881 | 912 | co | 21 | 0.7 | co | 28 | 7.5 | sp | 14 | 3.1 |
| 4 | 8 | 382 | 496 | 17686 | 2597 | co | 31 | 7.5 | co | 46 | 1121. | sp | 22 | 125. |
| 4 | 10 | 842 | 1002 | 79822 | 5953 | co | 46 | 67. | rl | * | * | tl | 12 | 1914. |
| 4 | 12 | 1626 | 1821 | 286441 | 11832 | co | 66 | 496. | m | * | * | sk | * | * |
| 6 | 4 | 80 | 211 | 1240 | 800 | co | 19 | 0.6 | co | 21 | 1.0 | co | 11 | 0.6 |
| 6 | 6 | 422 | 925 | 20539 | 5336 | co | 26 | 35. | co | 40 | 1347. | sp | 22 | 306. |
| 6 | 8 | 1514 | 3004 | 215440 | 22409 | co | 37 | 1276. | m | * | * | rl | * | * |
| 8 | 4 | 138 | 496 | 3412 | 2117 | co | 23 | 3.8 | co | 24 | 6.0 | co | 12 | 4.4 |
| 8 | 6 | 938 | 3004 | 89008 | 20825 | co | 38 | 1002. | rl | * | * | rl | * | * |
| 10 | 4 | 212 | 1002 | 7657 | 4633 | co | 26 | 27. | co | 28 | 54. | co | 12 | 25. |
| 12 | 4 | 302 | 1821 | 15003 | 8920 | co | 34 | 162. | co | 32 | 337. | co | 18 | 174. |
| 14 | 4 | 408 | 3061 | 26686 | 15662 | co | 37 | 652. | co | 37 | 1765. | sp | 16 | 626. |

Table 6.11: Shape constrained regression formulation and solver statistics. Note $\nu = \bar{\nu}$, $p = 0$, $\bar{p} = q - n - 1$, $\bar{q} = \bar{n} + 1$.

we propose NFs and generate instances of a wide range of sizes. For these instances, the EFs are often orders of magnitude larger than the NFs, and the EFs often have larger barrier parameters. We demonstrate significant computational advantages from solving the NFs with Hypatia compared to solving the EFs with either Hypatia or MOSEK 9, especially in terms of solve time and memory usage.

As we observe, NFs can be faster and less memory-intensive to generate using modeling tools such as JuMP, and they are typically more convenient to write and interpret conic certificates for. In deciding whether to formulate an NF or an EF, it can be helpful to examine our summary in Table 5.1 of computational properties for NFs and EFs of exotic cone constraints. For spectral and nuclear norm constraints, when the matrix ($W \in \mathbb{R}^{d_1 \times d_2}$) has many more columns than rows ($d_2 \gg d_1$), the dimensions and barrier parameter look relatively more favorable for the NF. For SOS and SOS matrix constraints, the dimensions grow much more slowly for the NF as the polynomial degree increases. Sometimes the modeler has to choose between different EFs. For our matrix completion problem and experiment design root-determinant variant, we compare two EFs for the geometric mean cone and find that Hypatia performs better with the exponential cone EF (EF-exp) and MOSEK performs better with the second order cone EF (EF-sec).

Our results clearly suggest that when there exists an NF that is significantly smaller than any EF, it is probably worth trying to solve the NF with Hypatia. If the NF uses a primitive proper cone not already defined in Hypatia, and the necessary LHSCB oracles (see section 4) for the cone or its dual cone are known and efficient to compute, the user can add support for the cone. If the user does not know an LHSCB for the cone or its dual cone, they can try using techniques such as those in

[27], which we use to derive LHSCBs for several new cones in section 4.

## Appendix A. Self-concordance of new barriers.

LEMMA A.1 (Logarithm cone LHSCB). *An LHSCB for $\mathcal{K}_{\log} = \mathrm{cl}\{(u, v, w) \in \mathbb{R} \times \mathbb{R}_{>}^{1+d} : u \leq \sum_{i \in \llbracket d \rrbracket} v \log(\frac{w_i}{v})\}$ is* (4.11).

*Proof.* We can write $\mathcal{K}_{\log} = \{(u, v, w) \in \mathbb{R}^{2+d} : (u, v, \mathrm{vec}(\mathrm{Diag}(w))) \in \mathcal{K}_{\mathrm{logdet}}\}$. So by [27, Proposition 5.1.1], (4.11) is a $(2 + d)$-self-concordant barrier for $\mathcal{K}_{\log}$. $\quad\square$

Our proposed LHSCBs for the log-determinant and root-determinant cones in (4.20) and (4.22) are clearly logarithmically homogeneous barriers by [27, Definition 2.3.2], so in Propositions A.2 and A.3 we prove self-concordance. For convenience, given $w \in \mathrm{int}(\mathcal{K}_{\succeq}) \subset \mathbb{R}^{\mathrm{sd}(d)}$, $z \in \mathbb{R}^{\mathrm{sd}(d)}$, we let $W = \mathrm{mat}(w) \in \mathbb{S}_{\succ}^d$, $Z = \mathrm{mat}(z) \in \mathbb{S}^d$. We also let $\lambda \in \mathbb{R}^d$ be the eigenvalues of the matrix $W^{-1/2} Z W^{-1/2} \in \mathbb{S}^d$, and let $t_k = \mathrm{tr}((W^{-1}Z)^k) = \sum_{i \in \llbracket d \rrbracket} \lambda_i^k, \forall k \in \llbracket 3 \rrbracket$.

PROPOSITION A.2 (Log-determinant cone LHSCB). *Consider the domain $\Gamma = \mathbb{R} \times \mathbb{R}_{\geq} \times \mathcal{K}_{\succeq} \subset \mathbb{R}^{2+\mathrm{sd}(d)}$. An LHSCB for $\mathcal{K}_{\mathrm{logdet}} = \mathrm{cl}\{(u, v, w) \in \mathrm{int}(\Gamma) : u \leq v \operatorname{logdet}(\frac{W}{v})\}$ is $f(u, v, w) = -\log(v) - \operatorname{logdet}(W) - \log(v \operatorname{logdet}(\frac{W}{v}) - u)$.*

*Proof.* Let $p = (u, v, w) \in \mathrm{int}(\Gamma)$. Let $g : \mathrm{int}(\Gamma) \mapsto \mathbb{R}$ be the function $g(p) = v \operatorname{logdet}(\frac{W}{v}) - u$. Given any direction $h = (x, y, z) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^{\mathrm{sd}(d)}$, the second and third directional derivatives of $g$ at $p$ are:

(A.1a) $\qquad \nabla^2 g(p)[h, h] = -v t_2 - 2 y t_1 + d v^{-1} y^2 = -v^{-1} \sum_{i \in \llbracket d \rrbracket} (v \lambda_i - y)^2,$

(A.1b) $\quad \nabla^3 g(p)[h, h, h] = 2 v t_3 - 3 y t_2 + d v^{-2} y^3 = v^{-2} \sum_{i \in \llbracket d \rrbracket} (v \lambda_i - y)^2 (2 v \lambda_i + y).$

Following [27, Definition 5.1.1], we claim that $g$ is $(\mathbb{R}_{\geq}, 1)$-compatible with the domain $\Gamma$. First, we note that $g$ is (i) $C^3$ smooth on $\mathrm{int}(\Gamma)$ and (ii) concave with respect to $\mathbb{R}_{\geq}$. Now it remains to show that (iii):

(A.2) $\qquad\qquad \nabla^3 g(p)[h, h, h] \leq -3 \nabla^2 g(p)[h, h] \qquad \forall p \in \mathrm{int}(\Gamma),$

for all $h = (x, y, z)$ such that $v \pm y \geq 0$, $W \pm Z \in \mathbb{S}_{\succeq}^d$, i.e. for all $h$ satisfying $|y| \leq v, |\lambda_i| \leq 1$. Since the directional derivatives in (A.1) decouple by each index $i$ in the sums, it is sufficient for (A.2) to hold separately for each $i \in \llbracket d \rrbracket$, which is true because $2 v \lambda_i + y \leq 3 v$ for all $y \leq v, \lambda_i \leq 1$. Now by [27, Proposition 5.1.7], $f$ is a $(2 + d)$-self-concordant barrier for $\mathcal{K}_{\mathrm{logdet}}$. $\quad\square$

The following proposition and its proof was suggested by A. Nemirovski in a personal communication [24].

PROPOSITION A.3 (Root-determinant cone LHSCB). *Consider the domain $\Gamma = \mathbb{R} \times \mathcal{K}_{\succeq} \subset \mathbb{R}^{1+\mathrm{sd}(d)}$. An LHSCB for $\mathcal{K}_{\mathrm{rtdet}} = \{(u, w) \in \Gamma, u^d \leq \det(W)\}$ is $f(u, w) = -(5/3)^2 (\operatorname{logdet}(W) + \log((\det(W))^{1/d} - u))$.*

*Proof.* Let $\mu = d^{-1} \sum_{i \in \llbracket d \rrbracket} \lambda_i$ be the mean of $\lambda$ and $\delta_i = \lambda_i - \mu, \forall \in \llbracket d \rrbracket$. Then:

(A.3) $\qquad t_1 = d\mu, \quad t_2 = d\mu^2 + \sum_{i \in \llbracket d \rrbracket} \delta_i^2, \quad t_3 = d\mu^3 + 3\mu \sum_{i \in \llbracket d \rrbracket} (\delta_i^2 + \delta_i^3).$

Let $p = (u, w) \in \mathrm{int}(\Gamma)$. Let $g : \mathrm{int}(\Gamma) \mapsto \mathbb{R}$ be the function $g(p) = (\det(W))^{1/d} - u$, and let $\phi = d^{-1} (\det(W))^{1/d} > 0$. Given any direction $h = (x, z) \in \mathbb{R} \times \mathbb{R}^{\mathrm{sd}(d)}$, the second and third directional derivatives of $g$ at $p$ are:

(A.4a) $\qquad \nabla^2 g(p)[h, h] = -\phi(t_2 - d^{-1} t_1^2) = -\phi \sum_{i \in \llbracket d \rrbracket} \delta_i^2,$

(A.4b) $\qquad \nabla^3 g(p)[h, h, h] = -\phi(-d^{-2} t_1^3 + 3 d^{-1} t_1 t_2 - 2 t_3) = \phi \sum_{i \in \llbracket d \rrbracket} (3\mu \delta_i^2 - \delta_i^3).$

Following [27, Definition 5.1.1], we claim that $g$ is $(\mathbb{R}_{\geq}, \beta)$-compatible with the domain $\Gamma$, where $\beta = 5/3$. First, we note that $g$ is (i) $C^3$ smooth on $\mathrm{int}(\Gamma)$ and (ii) concave with respect to $\mathbb{R}_{\geq}$. Now it remains to show that (iii):

$$\text{(A.5)} \qquad \nabla^3 g(p)[h, h, h] \leq -3\beta \nabla^2 g(p)[h, h] \qquad \forall p \in \mathrm{int}(\Gamma),$$

for all $h = (x, z)$ such that $W \pm Z \in \mathbb{S}_{\succeq}^d$, i.e. that:

$$\text{(A.6)} \qquad \sum_{i \in [\![d]\!]} (3\mu \delta_i^2 - \delta_i^3) \leq 3\beta \sum_{i \in [\![d]\!]} \delta_i^2,$$

since $\phi > 0$. Note that $W \pm Z \in \mathbb{S}_{\succeq}^d$ implies $|\lambda_i| \leq 1$, $|\mu| \leq 1$, and $|\delta_i| \leq 2$. Since the terms in the sums of (A.6) decouple by each index $i$, it is sufficient to show that inequality holds separately for each $i \in [\![d]\!]$. It can be verified this holds for $\beta = 5/3$. Now by [27, Proposition 5.1.7], $f$ is a $\beta^2(1 + d)$-self-concordant barrier for $\mathcal{K}_{\mathrm{rtdet}}$. $\quad\square$

## REFERENCES

[1] A. Agrawal, S. Diamond, and S. Boyd, *Disciplined geometric programming*, Optimization Letters, 13 (2019), pp. 961–976.

[2] E. D. Andersen, C. Roos, and T. Terlaky, *On implementing a primal-dual interior-point method for conic quadratic optimization*, Mathematical Programming, 95 (2003), pp. 249–277.

[3] M. Andersen, J. Dahl, Z. Liu, L. Vandenberghe, S. Sra, S. Nowozin, and S. Wright, *Interior-point methods for large-scale cone programming*, Optimization for Machine Learning, 5583 (2011).

[4] M. S. Andersen, J. Dahl, and L. Vandenberghe, *Implementation of nonsymmetric interior-point methods for linear optimization over sparse matrix cones*, Mathematical Programming Computation, 2 (2010), pp. 167–201.

[5] M. S. Andersen, J. Dahl, and L. Vandenberghe, *Logarithmic barriers for sparse matrix cones*, Optimization Methods and Software, 28 (2013), pp. 396–423.

[6] A. Ben-Tal and A. Nemirovski, *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*, vol. 2, SIAM, 2001.

[7] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, *Julia: A fresh approach to numerical computing*, SIAM Review, 59 (2017), pp. 65–98.

[8] B. Borchers, *CSDP, a C library for semidefinite programming*, Optimization Methods and Software, 11 (1999), pp. 613–623.

[9] S. Boyd and L. Vandenberghe, *Convex optimization*, Cambridge University Press, 2004.

[10] S. Burer, *Semidefinite programming in the space of partial positive semidefinite matrices*, SIAM Journal on Optimization, 14 (2003), pp. 139–172.

[11] C. Coey, M. Lubin, and J. P. Vielma, *Outer approximation with conic certificates for mixed-integer convex problems*, Mathematical Programming Computation, (2020), pp. 1–45.

[12] J. Dahl and E. D. Andersen, *A primal-dual interior-point algorithm for nonsymmetric exponential-cone optimization*, (2019), https://docs.mosek.com/whitepapers/expcone.pdf.

[13] S. Diamond and S. Boyd, *CVXPY: a Python-embedded modeling language for convex optimization*, The Journal of Machine Learning Research, 17 (2016), pp. 2909–2913.

[14] A. Domahidi, E. Chu, and S. Boyd, *ECOS: an SOCP solver for embedded systems*, in 2013 European Control Conference (ECC), IEEE, 2013, pp. 3071–3076.

[15] I. Dunning, J. Huchette, and M. Lubin, *JuMP: a modeling language for mathematical optimization*, SIAM Review, 59 (2017), pp. 295–320.

[16] M. Grant and S. Boyd, *CVX: MATLAB software for disciplined convex programming, version 2.1*, 2014.

[17] O. Güler, *Barrier functions in interior point methods*, Mathematics of Operations Research, 21 (1996), pp. 860–885.

[18] G. Hall, *Engineering and business applications of sum of squares polynomials*, arXiv preprint arXiv:1906.07961, (2019).

[19] M. Karimi and L. Tunçel, *Primal-dual interior-point methods for domain-driven formulations: algorithms*, arXiv preprint arXiv:1804.06925, (2018).

[20] B. Legat, O. Dowson, J. D. Garcia, and M. Lubin, *MathOptInterface: a data structure for mathematical optimization problems*, arXiv preprint arXiv:2002.03447, (2020).

[21] M. Lubin, E. Yamangil, R. Bent, and J. P. Vielma, *Extended formulations in mixed-integer convex programming*, in Integer Programming and Combinatorial Optimization: 18th International Conference, IPCO 2016, Liège, Belgium, June 1-3, 2016, Proceedings, Q. Louveaux and M. Skutella, eds., Springer International Publishing, 2016, pp. 102–113, https://doi.org/10.1007/978-3-319-33461-5_9.

[22] MOSEK ApS, *Modeling Cookbook revision 3.2.1*, 2020, https://docs.mosek.com/modeling-cookbook/index.html.

[23] MOSEK ApS, *MOSEK fusion API for Python*, 2020, https://docs.mosek.com/9.1/pythonfusion/index.html.

[24] A. Nemirovski. personal communication, 2019.

[25] Y. Nesterov, *Constructing self-concordant barriers for convex cones*, CORE discussion paper, (2006).

[26] Y. Nesterov, *Towards non-symmetric conic optimization*, Optimization Methods and Software, 27 (2012), pp. 893–917.

[27] Y. Nesterov and A. Nemirovskii, *Interior-point polynomial algorithms in convex programming*, Studies in Applied Mathematics, Society for Industrial and Applied Mathematics, 1994.

[28] Y. E. Nesterov and M. J. Todd, *Self-scaled barriers and interior-point methods for convex programming*, Mathematics of Operations Research, 22 (1997), pp. 1–42.

[29] Y. E. Nesterov, M. J. Todd, and Y. Ye, *Infeasible-start primal-dual methods and infeasibility detectors for nonlinear programming problems*, tech. report, Cornell University Operations Research and Industrial Engineering, 1996.

[30] B. O'Donoghue, E. Chu, N. Parikh, and S. Boyd, *Conic optimization via operator splitting and homogeneous self-dual embedding*, Journal of Optimization Theory and Applications, 169 (2016), pp. 1042–1068.

[31] D. Papp and F. Alizadeh, *Shape-constrained estimation using nonnegative splines*, Journal of Computational and Graphical Statistics, 23 (2014), pp. 211–231.

[32] D. Papp and S. Yildiz, *On "A homogeneous interior-point algorithm for non-symmetric convex conic optimization"*, arXiv preprint arXiv:1712.00492, (2017).

[33] D. Papp and S. Yildiz, *Sum-of-squares optimization without semidefinite programming*, SIAM Journal on Optimization, 29 (2019), pp. 822–851.

[34] F. Permenter, H. A. Friberg, and E. D. Andersen, *Solving conic optimization problems via self-dual embedding and facial reduction: a unified approach*, SIAM Journal on Optimization, 27 (2017), pp. 1257–1282.

[35] B. Recht, M. Fazel, and P. A. Parrilo, *Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization*, SIAM Review, 52 (2010), pp. 471–501.

[36] S. A. Serrano, *Algorithms for unsymmetric cone optimization and an implementation for problems with the exponential cone*, PhD thesis, Stanford University, 2015.

[37] A. Skajaa and Y. Ye, *A homogeneous interior-point algorithm for nonsymmetric convex conic optimization*, Mathematical Programming, 150 (2015), pp. 391–422.

[38] M. Udell, K. Mohan, D. Zeng, J. Hong, S. Diamond, and S. Boyd, *Convex optimization in Julia*, in Proceedings of the 1st First Workshop for High Performance Technical Computing in Dynamic Languages, IEEE Press, 2014, pp. 18–28.

[39] L. Vandenberghe, *The CVXOPT linear and quadratic cone program solvers*, Online: http://cvxopt. org/documentation/coneprog. pdf, (2010).

[40] X. Xu, P.-F. Hung, and Y. Ye, *A simplified homogeneous and self-dual linear programming algorithm and its implementation*, Annals of Operations Research, 62 (1996), pp. 151–171.

[41] M. Yamashita, K. Fujisawa, and M. Kojima, *Implementation and evaluation of SDPA 6.0 (semidefinite programming algorithm 6.0)*, Optimization Methods and Software, 18 (2003), pp. 491–505.

[42] J. Yang, L. Luo, J. Qian, Y. Tai, F. Zhang, and Y. Xu, *Nuclear norm based matrix regression with applications to face recognition with occlusion and illumination changes*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 39 (2016), pp. 156–171.