Bernoulli **7**(1), 2001, 79–97

Generalized Neyman–Pearson lemma via convex duality

JAKŠA CVITANIĆ[1] and IOANNIS KARATZAS[2]

[1]*Department of Mathematics, University of Southern California, Los Angeles, CA 90089. E-mail: cvitanic@math.usc.edu*
[2]*Departments of Mathematics and Statistics, Columbia University, New York, NY 10027. E-mail: ik@math.columbia.edu*

We extend the classical Neyman–Pearson theory for testing composite hypotheses versus composite alternatives, using a convex duality approach, first employed by Witting. Results of Aubin and Ekeland from non-smooth convex analysis are used, along with a theorem of Komlós, in order to establish the existence of a max-min optimal test in considerable generality, and to investigate its properties. The theory is illustrated on representative examples involving Gaussian measures on Euclidean and Wiener space.

Keywords: hypothesis testing; Komlós theorem; non-smooth convex analysis; normal cones; optimal generalized tests; saddle-points; stochastic games; subdifferentials

1. Introduction

On a measurable space $(\Omega, \mathscr{F})$, suppose that we are given two probability measures Q ('hypothesis') and P ('alternative'), and that we want to discriminate between them. We can try to do this in terms of a (pure) *test*, that is, a random variable $X : \Omega \to \{0, 1\}$, *which rejects* Q *on the event* $\{X = 1\}$. With this interpretation, $Q(X = 1)$ is the probability of rejecting Q when it is true (probability of Type I error), whereas $P(X = 0) = 1 - P(X = 1)$ is the probability of accepting Q when it is false (probability of Type II error).

Ideally, one would like to minimize these error probabilities simultaneously, but typically this will not be possible: a more sensitive radar decreases the chance of letting enemy aircraft go undetected, but also makes false alarms more likely. The next best thing is then to fix a certain number $0 < \alpha < 1$ (say $\alpha = 1\%$ or $\alpha = 5\%$), and try to

$$\text{maximize } P(X = 1), \qquad \text{subject to } Q(X = 1) \leq \alpha. \tag{1.1}$$

In other words, one tries to find a test that minimizes the probability of Type II error, among all tests that keep the probability of Type I error below the given acceptable *significance level* $\alpha \in (0, 1)$. This is the tack taken by the classical Neyman–Pearson theory of hypothesis testing; see, for instance, Lehmann (1986), Ferguson (1967) or Witting (1985).

The basic results of this theory are very well known. Suppose that μ is a third probability measure with

$$P \ll \mu, \qquad Q \ll \mu \tag{1.2}$$

1350–7265

(for instance, $\mu = (P + Q)/2$), and set

$$G := \frac{\mathrm{d}P}{\mathrm{d}\mu}, \qquad H := \frac{\mathrm{d}Q}{\mathrm{d}\mu}. \tag{1.3}$$

Then the problem of (1.1) has a solution, namely

$$\hat{X} = \mathbf{1}_{\{\hat{z}H < G\}}, \tag{1.4}$$

provided that

$$Q(\hat{z}H < G) = \alpha \qquad \text{for some} \qquad 0 < \hat{z} < \infty. \tag{1.5}$$

In other words, the test $\hat{X}$ of (1.4) rejects the hypothesis if and only if the 'likelihood ratio' $G/H = (\mathrm{d}P/\mathrm{d}\mu)/(\mathrm{d}Q/\mathrm{d}\mu)$ is sufficiently large.

When a number $\hat{z}$ with the properties (1.5) cannot be found, one has to consider *randomized tests*, that is, random variables $X : \Omega \to [0, 1]$. The new interpretation is that, if the outcome $\omega \in \Omega$ is observed, then the hypothesis Q is rejected (accepted) with probability $X(\omega)$ $(1 - X(\omega))$, independently of everything else. Thus,

$$\mathrm{E}^P(X) = \int X(\omega)P(\mathrm{d}\omega) \tag{1.6}$$

is the *power of the randomized test* X, that is, the probability of rejecting the hypothesis Q when it is false; and

$$\mathrm{E}^Q(X) = \int X(\omega)Q(\mathrm{d}\omega) \tag{1.7}$$

is the probability of Type I error for the randomized test X (i.e., of rejecting Q when it is true). By analogy with (1.1), one seeks a randomized test $\hat{X}$ which will

$$\text{maximize } \mathrm{E}^P(X), \text{ over all randomized tests } X : \Omega \to [0, 1] \text{ with } \mathrm{E}^Q(X) \leqslant \alpha. \tag{1.8}$$

The advantage of this 'randomized' formulation is that (1.8) has a solution for *any* given significance level $\alpha \in (0, 1)$. In particular, the supremum

$$\sup_{X \in \mathscr{X}_\alpha} \mathrm{E}^P(X), \qquad \text{with } \mathscr{X}_\alpha := \{X : \Omega \to [0, 1];\ \mathrm{E}^Q(X) \leqslant \alpha\}, \tag{1.9}$$

is attained by the randomized test

$$\hat{X} = 1_{\{\hat{z}H < G\}} + b \cdot \mathbf{1}_{\{\hat{z}H = G\}} \tag{1.10}$$

in $\mathscr{X}_\alpha$, where we have set (with the convention $0/0 = 0$):

$$\hat{z} := \inf\{u \geqslant 0;\ Q(uH < G) \leqslant \alpha\}, \tag{1.11}$$

$$b := \frac{\alpha - Q(\hat{z}H < G)}{Q(\hat{z}H = G)} = \frac{\alpha - Q(\hat{z}H < G)}{Q(\hat{z}H \leqslant G) - Q(\hat{z}H < G)} \in [0, 1]. \tag{1.12}$$

2. Composite hypotheses and alternatives

Let us suppose now that, on the measurable space $(\Omega, \mathscr{F})$, we have an entire *family* $\mathbf{Q}$ of probability measures (composite 'hypothesis'), which we want to discriminate against another family $\mathbf{P}$ of probability measures (composite 'alternative'). By analogy with (1.2), we asssume

$$\mathbf{P} \cap \mathbf{Q} = \varnothing, \tag{2.1}$$

$$P \ll \mu, \quad Q \ll \mu, \qquad \forall\, P \in \mathbf{P}, \quad \forall\, Q \in \mathbf{Q}, \tag{2.2}$$

for some probability measure μ, and set

$$G_P := \frac{\mathrm{d}P}{\mathrm{d}\mu}\ (P \in \mathbf{P}), \qquad H_Q := \frac{\mathrm{d}Q}{\mathrm{d}\mu}(Q \in \mathbf{Q}), \tag{2.3}$$

$$\mathscr{X}_\alpha := \{X : \Omega \to [0, 1];\ \mathrm{E}^Q(X) \leqslant \alpha, \forall Q \in \mathbf{Q}\} \tag{2.4}$$

as in (1.3), (1.9). Since we have now an entire family $\mathbf{P}$ of alternatives, we shall replace (1.9) by the *max-min criterion*

$$\underline{V}(\alpha) := \sup_{X \in \mathscr{X}_\alpha} \left(\inf_{P \in \mathbf{P}} \mathrm{E}^P(X) \right). \tag{2.5}$$

In other words, we shall look for a randomized test $\hat{X}$ that maximizes the *smallest power*

$$\gamma(X) := \inf_{P \in \mathbf{P}} \mathrm{E}^P(X) \tag{2.6}$$

attainable through measures of the family $\mathbf{P}$, over all randomized tests X of *size*

$$s(X) := \sup_{Q \in \mathbf{Q}} \mathrm{E}^Q(X) \leqslant \alpha. \tag{2.7}$$

Definition 2.1. *If such a randomized test $\hat{X} \in \mathscr{X}_\alpha$ exists, it will be called* max-min-optimal *for testing the (composite) hypothesis* $\mathbf{Q}$ *against the (composite) alternative* $\mathbf{P}$*, at the given level of significance* $\alpha \in (0, 1)$.

Under appropriate conditions on the family $\mathbf{P}$ of alternatives, we shall see in the next section that an optimal max-min randomized test exists and has a form reminiscent of (1.10), namely

$$\hat{X} = 1_{\{\hat{z}\hat{H} < \hat{G}\}} + B \cdot \mathbf{1}_{\{\hat{z}\hat{H} = \hat{G}\}}. \tag{2.8}$$

Here B is a *random variable* with values in the interval $[0, 1]$; the random variable $\hat{G}$ is of the form $G_{\hat{P}} = \mathrm{d}\hat{P}/\mathrm{d}\mu$ of (2.3) for some $\hat{P} \in \mathbf{P}$; the random variable $\hat{H}$ is chosen from a suitable family that contains the convex hull

$$Co(\mathrm{H}; \mathbf{Q}) := \{\lambda H_{Q_1} + (1-\lambda) H_{Q_2};\ Q_1 \in \mathbf{Q}, Q_2 \in \mathbf{Q}, 0 \leqslant \lambda \leqslant 1\} \tag{2.9}$$

of $\{H_Q\}_{Q \in \mathbf{Q}}$; and $\hat{z}$ is a suitable positive number.

As we shall see, it is no longer possible to compute these quantities by means of formulae as explicit as (1.11) and (1.12), which are valid in the simple hypothesis versus simple alternative case. However, methods of *non-smooth convex analysis* and *duality theory* provide both the existence of these quantities and algorithms that can lead to their computation, as illustrated by several examples in Section 5. The main result of the present paper, namely Theorem 4.1 (and its Corollaries 4.1 and 4.2), shows that the associated dual problem *always* (that is, even with uncountably many hypotheses and/or alternatives) has a solution; that there is *never* a duality gap; and that the optimal test *always* has the 0–1 representation (2.8).

The idea of using convex duality methods in hypothesis testing is not new; it goes back to the paper by Krafft and Witting (1967), and is developed to a considerable extent in Chapter 2 of the treatise by Witting (1985), particularly propositions 2.80 and 2.81 on pp. 267–274. See also the papers by Lehmann (1952), Baumann (1968) and Huber and Strassen (1973), as well as Vajda (1989, pp. 361–362), for related results. Baumann (1968) proves the existence of the max-min-optimal test using general duality results from the theory of linear programming, as well as weak-compactness arguments. We provide a different, self-contained proof for existence, using an almost sure convergence argument based on the theorem of Komlós (1967). In addition, our approach enables us to show that the optimal test is *always* of the Neyman–Pearson form (2.8), thereby giving a characterization potentially useful in finding algorithms for computing the optimal test. To the best of our knowledge, the characterization (2.8) has hitherto been known only under stronger conditions on the null and the alternative hypotheses. We obtain it here in a very general setting, using infinite-dimensional non-smooth convex optimization results of Aubin and Ekeland (1984), apparently not employed previously in the theory of hypothesis testing. Our own inspiration came from Heath (1993), who used the Neyman–Pearson lemma as a tool for solving a stochastic control problem that can also be treated by methods of convex duality; for related work along these lines, see Karatzas (1997), as well as Cvitanić and Karatzas (1999), Cvitanić (2000), Föllmer and Leukert (1999; 2000) and Spivak (1998) for similar problems arising in the context of mathematical finance.

3. Results: Analysis

Let us begin the statement of results by introducing the set of random variables

$$\mathscr{H} := \{H \in L^1(\mu)/H \geqslant 0,\ \mu\text{-a.e. and } \mathrm{E}^{\mu}(HX) \leqslant \alpha,\ \forall X \in \mathscr{X}_\alpha\}. \tag{3.1}$$

As is relatively straightforward to check (cf. Section 6), this set is convex, bounded in $L^1(\mu)$, closed under μ-almost everywhere (a.e.) convergence, and contains the convex hull of (2.9), namely

$$Co(H; \mathbf{Q}) \subseteq \mathscr{H}. \tag{3.2}$$

In a similar spirit, we shall impose the following assumption throughout.

Assumption 3.1. *The set of densities*

$$\mathscr{G} := \{G_P\}_{P \in \mathbf{P}} \tag{3.3}$$

is convex and closed under μ-a.e. convergence. (*From* (2.3), *the convexity of* $\mathscr{G}$ *is equivalent to the convexity of the family* $\mathbf{P}$ *of alternatives*.)

The set $\mathscr{G}$ of (3.3) is obviously bounded in $L^1(\mu)$, since $\mathrm{E}^\mu(G_P) = P(\Omega) = 1$ for every $P \in \mathbf{P}$. We shall comment in Remark 4.1 below on the necessity of imposing Assumption 3.1 and of considering the class $\mathscr{H}$ of (3.1) as we do. Now the key observation is that, for arbitrary $G \in \mathscr{G}$ and $H \in \mathscr{H}$, we have $G = \mathrm{d}P/\mathrm{d}\mu$ for some $P \in \mathbf{P}$, and thus

$$\begin{aligned} \mathrm{E}^P(X) &= \mathrm{E}^\mu(GX) = \mathrm{E}^\mu[X(G - zH)] + z \cdot \mathrm{E}^\mu(HX) \\ &\leqslant \mathrm{E}^\mu(G - zH)^+ + \alpha z; \qquad \forall z > 0, \forall X \in \mathscr{X}_\alpha \end{aligned} \tag{3.4}$$

from (3.1) and $0 \leqslant X \leqslant 1$ in (1.9). Furthermore, we have equality in (3.4) for some $\hat{G} \in \mathscr{G}$, $\hat{H} \in \mathscr{H}$, $\hat{z} \in (0, \infty)$ and $\hat{X} \in \mathscr{X}_\alpha$, *if and only if* the conditions

$$\mathrm{E}^\mu(\hat{H}\hat{X}) = \alpha, \tag{3.5}$$

$$\hat{X} = \mathbf{1}_{\{\hat{z}\hat{H} < \hat{G}\}} + B \cdot \mathbf{1}_{\{\hat{z}\hat{H} = \hat{G}\}}, \qquad \mu\text{-a.e.}, \tag{2.8}$$

both hold, for some random variable $B: \Omega \to [0, 1]$; and in this case, with $\hat{P} := \int \hat{G} \mathrm{d}\mu \in \mathbf{P}$, we have

$$\mathrm{E}^{\hat{P}}(\hat{X}) = \mathrm{E}^\mu(\hat{G}\hat{X}) = \alpha\hat{z} + \mathrm{E}^\mu(\hat{G} - \hat{z}\hat{H})^+. \tag{3.6}$$

Proposition 3.1. *Suppose that there exists a quadruple* $(\hat{G}, \hat{H}, \hat{z}, \hat{X}) \in (\mathscr{G} \times \mathscr{H} \times (0, \infty) \times \mathscr{X}_\alpha)$ *that satisfies* (3.5), (2.8) *and*

$$\mathrm{E}^\mu[\hat{X}(\hat{G} - G)] \leqslant 0, \qquad \forall G \in \mathscr{G}. \tag{3.7}$$

Then we have

$$\mathrm{E}^{\hat{P}}(X) \leqslant \mathrm{E}^{\hat{P}}(\hat{X}) \leqslant \mathrm{E}^P(\hat{X}); \qquad \forall X \in \mathscr{X}_\alpha, \forall\, P \in \mathbf{P}. \tag{3.8}$$

In other words, the pair $(\hat{X}, \hat{P})$, *with* $\hat{P} = \int \hat{G}\, \mathrm{d}\mu$, *is then a saddle-point for the stochastic game with lower value* $\underline{V}(\alpha)$ *as in* (2.5), *and upper value*

$$\overline{V}(\alpha) := \inf_{P \in \mathbf{P}} \left(\sup_{X \in \mathscr{X}_\alpha} \mathrm{E}^P(X) \right), \tag{3.9}$$

namely

$$\underline{V}(\alpha) = \overline{V}(\alpha) = \mathrm{E}^{\hat{P}}(\hat{X}) = \int \hat{G}\hat{X}\, \mathrm{d}\mu. \tag{3.10}$$

We postpone to Section 6 the proofs of all the results in this paper.

Inequality (3.4) also points the way to a *duality approach*, which will lead eventually to the existence of a quadruple $(\hat{G}, \hat{H}, \hat{z}, \hat{X})$ with the properties (3.5)–(3.7) and (2.8), as follows. Let us introduce the value function

$$\tilde{V}(z) \equiv \tilde{V}(z; \alpha) := \inf_{(G,H) \in (\mathscr{G} \times \mathscr{H})} \mathrm{E}^\mu(G - zH)^+, \qquad 0 < z < \infty, \tag{3.11}$$

of an *auxiliary dual problem*, and observe from (3.4) the inequality

$$\overline{V}(\alpha) \leqslant V_*(\alpha), \tag{3.12}$$

where we have set

$$V_*(\alpha) := \inf_{\substack{z>0\\(G,H)\in(\mathscr{G}\times\mathscr{H})}} [\alpha z + \mathrm{E}^\mu(G - zH)^+] = \inf_{z>0}[\alpha z + \tilde{V}(z)]. \tag{3.13}$$

Proposition 3.2. *Under the assumptions of Proposition* 3.1*, the following hold:*

(i) The pair $(\hat{G}, \hat{H})$ *attains the infimum in* (3.11) *with* $z = \hat{z}$.
(ii) The triple $(\hat{z}, \hat{G}, \hat{H})$ *attains the first infimum in* (3.13).
(iii) The number $\hat{z} \in (0, \infty)$ *attains the second infimum in* (3.13).
(iv) There is no 'duality gap' in (3.12)*; that is,* $V_*(\alpha) = \overline{V}(\alpha) = \underline{V}(\alpha) = \mathrm{E}^{\hat{P}}(\hat{X})$.

We shall show in the next section how to select the 'dual variables' $(\hat{z}, \hat{G}, \hat{H}) \in ((0, \infty) \times \mathscr{G} \times \mathscr{H})$ in such a way that the optimal 'primal variable' (generalized test) $\hat{X}$ is then given in the form (2.8).

4. Results: Synthesis

We can now follow the above reasoning in reverse, and try to obtain the existence of the quadruple $(\hat{G}, \hat{H}, \hat{z}, \hat{X})$ postulated in Proposition 3.1 by characterizing its constituent elements in terms of the properties of Proposition 3.2. This is done in Lemma 4.1–4.4 and in Theorem 4.1 below, using *non-smooth convex analysis* as our main tool; cf. Aubin & Ekeland (1984, Chapters 1–4).

Lemma 4.1. *The function* $\tilde{V}(\cdot)$ *of* (3.11) *is Lipschitz-continuous:*

$$|\tilde{V}(z_1) - \tilde{V}(z_2)| \leqslant |z_1 - z_2|, \qquad \forall\, 0 < z_1, z_2 < \infty. \tag{4.1}$$

Lemma 4.2. *For any given* $z \geqslant 0$, *there exists a pair* $(\hat{G}, \hat{H}) = (\hat{G}_z, \hat{H}_z) \in \mathscr{G} \times \mathscr{H}$ *that attains the infimum in* (3.11).

Lemma 4.3. *For any given* $\alpha \in (0, 1)$, *there exists a number* $\hat{z} = \hat{z}_\alpha > 0$ *that attains the infimum* $V_*(\alpha) = \inf_{z>0}[\alpha z + \tilde{V}(z)]$ *of* (3.13).

Lemma 4.4. *Under the norm* $\|(G, H, z)\| := \mathrm{E}^\mu(|G| + |H|) + |z|$, *the set*

$$\mathscr{M} := \{(G, H, z) \in \mathscr{L};\ G \in \mathscr{G},\ H \in \mathscr{H},\ z \geqslant 0\} \tag{4.2}$$

is a closed, convex subset of $\mathscr{L} := L^1(\mu) \times L^1(\mu) \times \mathbb{R}$. *Furthermore, the functional*

$$\mathscr{L} \ni (G, H, z) \mapsto \tilde{U}(G, H, z) := \alpha z + \mathrm{E}^\mu(G - H)^+ \in (0, \infty) \tag{4.3}$$

is proper, convex and lower-semicontinuous on $\mathscr{L}$*; and* $(G, H, z) \mapsto \tilde{U}(G, zH, z) =$

$\alpha z + \mathrm{E}^{\mu}(G - zH)^{+}$ attains its infimum over $\mathscr{M}$ at the triple $(\hat{G}, \hat{H}, \hat{z})$, with $\hat{z} \equiv \hat{z}_{\alpha}$ as in Lemma 4.3 and $(\hat{G}, \hat{H}) \equiv (\hat{G}_{\hat{z}}, \hat{H}_{\hat{z}})$ as in Lemma 4.2.

Let us now consider the dual $\mathscr{L}^{*} := L^{\infty}(\mu) \times L^{\infty}(\mu) \times \mathbb{R}$ of the space $\mathscr{L} = L^{1}(\mu) \times L^{1}(\mu) \times \mathbb{R}$, the set $\mathscr{M}_{*} := \{(G, zH, z); (G, H, z) \in \mathscr{M}\}$, the *normal cone*

$$N(\hat{G}, \hat{z}\hat{H}, \hat{z}) := \{(W, X, y) \in \mathscr{L}^{*};\ \mathrm{E}^{\mu}(\hat{G}W + \hat{z}\hat{H}X) + \hat{z}y \geqslant \mathrm{E}^{\mu}(GW + zHX) + zy,\ \forall (G, H, z) \in \mathscr{M}\} \tag{4.4}$$

to $\mathscr{M}_{*}$ at the point $(\hat{G}, \hat{z}\hat{H}, \hat{z}) \in \mathscr{M}_{*}$, and the *subdifferential* at this point

$$\partial\tilde{U}(\hat{G}, \hat{z}\hat{H}, \hat{z}) := \{(W, X, y) \in \mathscr{L}^{*};\ \tilde{U}(\hat{G}, \hat{z}\hat{H}, \hat{z}) - \tilde{U}(G, Y, z) \leqslant \mathrm{E}^{\mu}[W(\hat{G} - G) + X(\hat{z}\hat{H} - Y)] + y(\hat{z} - z),\ \forall (G, Y, z) \in \mathscr{L}\} \tag{4.5}$$

of the functional $\tilde{U}$ in (4.3). From Lemma 4.4 and non-smooth convex analysis, as in Corollary 4.6.3 of Aubin and Ekeland (1984), we know that

$$(\hat{G}, \hat{z}\hat{H}, \hat{z}) \in \mathscr{M}_{*} \qquad \text{is a solution of} \qquad (0, 0, 0) \in \partial\tilde{U}(\hat{G}, \hat{z}\hat{H}, \hat{z}) + N(\hat{G}, \hat{z}\hat{H}, \hat{z}).$$

In other words, *there exists a triple* $(\hat{W}, \hat{X}, \hat{y})$, *such that* $(\hat{W}, \hat{X}, \hat{y}) \in N(\hat{G}, \hat{z}\hat{H}, \hat{z})$ *and* $(-\hat{W}, -\hat{X}, -\hat{y}) \in \partial\tilde{U}(\hat{G}, \hat{z}\hat{H}, \hat{z})$, or equivalently

$$\mathrm{E}^{\mu}[\hat{W}(\hat{G} - G)] + \mathrm{E}^{\mu}[\hat{X}(\hat{z}\hat{H} - zH)] + \hat{y}(\hat{z} - z) \geqslant 0, \qquad \forall (G, H, z) \in \mathscr{M}, \tag{4.6}$$

$$\mathrm{E}^{\mu}[\hat{W}(G - \hat{G}) + \hat{X}(Y - \hat{z}\hat{H}) + (G - Y)^{+} - (\hat{G} - \hat{z}\hat{H})^{+}] \geqslant (\alpha + \hat{y})(\hat{z} - z), \qquad \forall (G, Y, z) \in \mathscr{L}. \tag{4.7}$$

Sending $z \to \pm\infty$, we observe that (4.7) can hold only if

$$\hat{y} = -\alpha. \tag{4.8}$$

Substituting $\hat{y} = -\alpha$ as well as $G = \hat{G}$, $H = \hat{H}$, $z = \hat{z} \pm \delta$ for $\delta > 0$ into (4.6), we obtain

$$\mathrm{E}^{\mu}(\hat{H}\hat{X}) = \alpha \tag{3.5a}$$

as postulated by (3.5). On the other hand, substituting $G = \hat{G}$, $z = \hat{z}$ into (4.6), we obtain

$$\mathrm{E}^{\mu}(H\hat{X}) \leqslant \mathrm{E}^{\mu}(\hat{H}\hat{X}) = \alpha, \qquad \forall\ H \in \mathscr{H}, \tag{3.5b}$$

in conjunction with (3.5a); and substituting $H = \hat{H}$, $z = \hat{z}$ into (4.6) leads to

$$\mathrm{E}^{\mu}[\hat{W}(\hat{G} - G)] \geqslant 0, \qquad \forall\ G \in \mathscr{G}. \tag{3.7$'$}$$

Theorem 4.1. *The random variable $\hat{X}$ belongs to $\mathscr{X}_{\alpha}$ and satisfies*

$$\hat{X} = -\hat{W} = \mathbf{1}_{\{\hat{z}\hat{H} < \hat{G}\}} + B \cdot \mathbf{1}_{\{\hat{z}\hat{H} = \hat{G}\}}, \qquad \mu\text{-a.e.}, \tag{4.9}$$

for a suitable random variable $B : \Omega \to [0, 1]$.

Corollary 4.1. *The stochastic game with lower* (*upper*) *value $\underline{V}(\alpha)$ $(\overline{V}(\alpha))$ as in* (2.5), (3.9) *has saddle-point $(\hat{X}, \hat{P})$ with $\hat{P} = \int \hat{G}\,\mathrm{d}\mu$, and value $\underline{V}(\alpha) = \overline{V}(\alpha) = \mathrm{E}^{\hat{P}}(\hat{X})$.*

This follows from Proposition 3.1, since the conditions (3.5) and (2.8) are satisfied (recall (4.9) and (3.5a)), and so is condition (3.7), because $(3.7)'$ and (4.9) imply

$$\mathrm{E}^{\mu}[\hat{X}(\hat{G}-G)] = \mathrm{E}^{\mu}[\hat{W}(G-\hat{G})] \leqslant 0, \qquad \forall G \in \mathscr{G}.$$

Corollary 4.2. *The randomized test $\hat{X} \in \mathscr{X}_\alpha$ is max-min optimal for testing the composite hypothesis* **Q** *against the composite alternative* **P** *in the sense of Definition* 2.1*; that is, it maximizes the smallest value of the power $\gamma(X) = \inf_{P \in \mathbf{P}} \mathrm{E}^P(X)$ attainable over alternatives $P \in \mathbf{P}$, against all generalized tests $X \in \mathscr{X}_\alpha$ of size $s(X) := \sup_{Q \in \mathbf{Q}} \mathrm{E}^Q(X) \leqslant \alpha$.*

Remark 4.1. If the convex set $\tilde{\mathscr{H}} := Co(\mathrm{H}; \mathbf{Q})$ of (2.9) is itself closed under μ-a.e. convergence, there is no need to introduce the larger set $\mathscr{H}$ of (3.1) and (3.2), since the auxiliary dual problem

$$\tilde{V}(z) := \inf_{(G,H) \in (\mathscr{G} \times \tilde{H})} \mathrm{E}^{\mu}(G - zH)^{+} \qquad (3.11)'$$

then has a solution in $\mathscr{G} \times \tilde{H}$.

Similarly, it is well known that an optimal test can often be found among ‘Bayesian tests’ (see Witting 1985, for example). More precisely, suppose that the set $\{H_Q;\, Q \in \mathbf{Q}\}$ can be represented in the form $\{H_\theta;\, \theta \in \Theta\}$ for some measurable space $\{\Theta, \mathscr{E}\}$, such that $(\theta, \omega) \mapsto H_\theta(\omega)$ is an $\mathscr{F} \otimes \mathscr{E}$-measurable function on $\Omega \times \Theta$. Let $\mathscr{S}$ be a set of probability measures (‘prior distributions’) on $\{\Theta, \mathscr{E}\}$, and write

$$\mathscr{H}_{\mathscr{S}} := \left\{ \int H_\theta \,\mathrm{d}S(\theta);\, S \in \mathscr{S} \right\}.$$

Again, if $\mathscr{H}_{\mathscr{S}}$ is convex and closed under μ-a.e. convergence, and if $\{H_Q;\, Q \in \mathbf{Q}\} \subseteq \mathscr{H}_{\mathscr{S}}$, we can work with the set $\mathscr{H}_{\mathscr{S}}$ instead of $\mathscr{H}$.

However, we *cannot* relax Assumption 3.1 on the set $\mathscr{G}$ of (3.3); in particular, we cannot replace $\mathscr{G}$ by the larger convex set

$$\tilde{\mathscr{G}} := \{G \in L^1(\mu);\, G \geqslant 0,\, \mu\text{-a.e.};\, \mathrm{E}^{\mu}(GX) \geqslant \inf_{P \in \mathbf{P}} \mathrm{E}^P(X),\, \forall X \in \mathscr{X}_\alpha\}. \qquad (3.3)'$$

One reason why this might not work, is that $\tilde{\mathscr{G}}$ does not have to be closed under μ-a.e. convergence, since Fatou’s lemma can then fail to produce the desired inequality in $(3.3)'$.

5. Examples

We present in this section a few representative examples involving hypothesis testing for Gaussian measures on Euclidean and Wiener spaces. Most, if not all, of the examples are probably well known; they are developed here only in so far as they allow us to illustrate the theory of Sections 3 and 4 in a transparent and direct way.

Let us consider first the case of the filtered measurable space $\{\Omega, \mathscr{F}\}$, $\mathbf{F} := \{\mathscr{F}(t)\}_{0 \leqslant t \leqslant 1}$, where $\Omega := C([0, 1]; \mathbb{R}^d)$ is the set of all continuous functions $\omega : [0, 1] \to \mathbb{R}^d$; $W(t, \omega) := \omega(t)$, $0 \leqslant t \leqslant 1$, is the coordinate mapping process;

$\mathscr{F}(t) := \sigma(W(s);\, 0 \leqslant s \leqslant t)$ and $\mathscr{F} = \mathscr{F}(1)$. Consider also the space Θ of $\mathbf{F}$-progressively measurable processes $\theta\colon [0, 1] \times \Omega \to K$ for some given compact, convex subset K of $\mathbb{R}^d$ with $0 \notin K$. We shall look at *Wiener measure* μ on $(\Omega, \mathscr{F})$, as well as at the family of measures $\{\mu_\theta,\, \theta \in \Theta\}$ given by

$$\frac{\mathrm{d}\mu_\theta}{\mathrm{d}\mu} = Z_\theta(1), \qquad \text{with} \qquad Z_\theta(t) := \exp\left\{\int_0^t \theta'(s)\,\mathrm{d}W(s) - \frac{1}{2}\int_0^t \|\theta(s)\|^2\,\mathrm{d}s\right\} \tag{5.1}$$

for $0 \leqslant t \leqslant 1$, $\theta(\cdot) \in \Theta$. From the Girsanov theorem (Karatzas and Shreve 1991, p. 191) we know that *under* μ_θ*, the process* $W_\theta(\cdot) := W(\cdot) - \int_0^{\cdot} \theta(s)\,\mathrm{d}s$ *is Brownian motion* (or, equivalently, that $W(\cdot)$ is Brownian motion with random drift $\theta(\cdot)$).

Example 5.1 *Simple hypothesis* $\mathbf{Q} = \{\mu\}$ *versus composite alternative* $\mathbf{P} = \{\mu_\theta,\, \theta \in \Theta\}$. In this case $H = 1$, $G_\theta := \mathrm{d}\mu_\theta/\mathrm{d}\mu = Z_\theta(1)$, and the set $\mathscr{G} := \{G_\theta\}_{\theta\in\Theta} = \{Z_\theta(1)\}_{\theta\in\Theta}$ is convex, bounded in $L^p(\mu)$ for every $p > 1$, and closed under μ-a.e. convergence (see Beneš 1971, pp. 463 and 469). Suppose that

$$\inf_{\vartheta\in K} \|\vartheta\| = \|\hat{\vartheta}\|, \qquad \text{for some} \qquad \hat{\vartheta} \in K. \tag{5.2}$$

Then it can be shown (see Section 6 for details) that the value function of the auxiliary dual problem in (3.11) is given by

$$\begin{aligned}\tilde{V}(z) &= \inf_{\theta\in\Theta} \mathrm{E}^\mu(Z_\theta(1) - z)^+ = \mathrm{E}^\mu(Z_{\hat{\vartheta}}(1) - z)^+ \\ &= \int_{\mathbb{R}^d} (\mathrm{e}^{\hat{\vartheta}\xi - \frac{1}{2}\|\hat{\vartheta}\|^2} - z)^+ \frac{\mathrm{e}^{-\frac{1}{2}\|\xi\|^2}}{(2\pi)^{d/2}}\,\mathrm{d}\xi.\end{aligned} \tag{5.3}$$

Denoting by $\hat{z}$ the unique number in $(0, \infty)$ where $z \mapsto \alpha z + \tilde{V}(z)$ attains its minimum, we obtain the optimal test of Theorem 4.1 in the (pure, non-randomized) form

$$\hat{X} = \mathbf{1}_{\{G_{\hat{\vartheta}} > \hat{z}\}} = \mathbf{1}_{\{\hat{\vartheta}' W(1) > \frac{1}{2}\|\hat{\vartheta}\|^2 + \log\hat{z}\}}. \tag{5.4}$$

This test rejects the hypothesis $Q = \mu$, if the inner product of $W(1)$ with the vector $\hat{\vartheta} \in K$ is sufficiently large. With $\Phi(x) \equiv (2\pi)^{-1/2}\int_{-\infty}^x \mathrm{e}^{-u^2/2}\,\mathrm{d}u$, we then have

$$\alpha = \mathrm{E}^\mu(\hat{X}) = \mu[\hat{\vartheta}' W(1) > \|\hat{\vartheta}\|^2/2 + \log\hat{z}] = 1 - \Phi\left(\frac{\log\hat{z} + \frac{1}{2}\|\hat{\vartheta}\|^2}{\|\hat{\vartheta}\|}\right),$$

and thus

$$\log\frac{1}{\hat{z}} = \frac{1}{2}\|\hat{\vartheta}\|^2 + \|\hat{\vartheta}\| \cdot \Phi^{-1}(\alpha). \tag{5.5}$$

Example 5.2 *Composite hypothesis* $\mathbf{Q} = \{\mu_\theta,\, \theta \in \Theta\}$ *versus simple alternative* $\mathbf{P} = \{\mu\}$. Here $G = 1$ and $H_\theta := \mathrm{d}\mu_\theta/\mathrm{d}\mu = Z_\theta(1)$, so that the set $\{H_\theta\}_{\theta\in\Theta}$ is convex and closed under μ-a.e. convergence (see Remark 4.1). It is shown in Section 6 that the auxiliary dual problem of (3.11) has value function

$$
\begin{aligned}
\tilde{V}(z) &= \inf_{\theta\in\Theta} \mathrm{E}^{\mu}(1 - zZ_{\theta}(1))^{+} = \inf_{\theta\in\Theta} \mathrm{E}^{\mu}\left(1 - z \cdot \mathrm{e}^{\int_0^1 \theta'(s)\,\mathrm{d}W(s) - \frac{1}{2}\int_0^1 \|\theta(s)\|^2\,\mathrm{d}s}\right)^{+} \\
&= \mathrm{E}^{\mu}(1 - zZ_{\hat{\vartheta}}(1))^{+} = \mathrm{E}^{\mu}\left(1 - z \cdot \mathrm{e}^{\hat{\vartheta}' W(1) - \frac{1}{2}\|\hat{\vartheta}\|^2}\right)^{+} \qquad (5.6) \\
&= \int_{\mathbb{R}^d} \left(1 - z \cdot \mathrm{e}^{\hat{\vartheta}'\xi - \frac{1}{2}\|\hat{\vartheta}\|^2}\right)^{+} \frac{\mathrm{e}^{-\frac{1}{2}\|\xi\|^2}}{(2\pi)^{d/2}}\,\mathrm{d}\xi
\end{aligned}
$$

under the assumption (5.2). With $\hat{z} = \arg\min_{0<z<\infty}[\alpha z + \tilde{V}(z)]$ the number of Lemma 4.3, we have now the optimal test of Theorem 4.1 in the (pure, non-randomized) form

$$\hat{X} = \mathbf{1}_{\{\hat{z} H_{\hat{\vartheta}} < 1\}} = \mathbf{1}_{\{\hat{\vartheta}' W(1) < \frac{1}{2}\|\hat{\vartheta}\|^2 - \log \hat{z}\}}. \qquad (5.7)$$

This test rejects the hypothesis, if the inner product of $W(1)$ with the vector $\hat{\vartheta}$ of (5.2) is sufficiently small. In particular, we have

$$
\begin{aligned}
\alpha &= \sup_{\theta\in\Theta} \mathrm{E}^{\mu_\theta}(\hat{X}) = \sup_{\theta\in\Theta} \mu_\theta[\hat{\vartheta}' W(1) < \|\hat{\vartheta}\|^2/2 - \log\hat{z}] \\
&= \sup_{\theta\in\Theta} \mu_\theta\left[\hat{\vartheta}'\left(W_\theta(1) + \int_0^1 \theta(s)\,\mathrm{d}s\right) < \|\hat{\vartheta}\|^2/2 - \log\hat{z}\right] \\
&= \mu_{\hat{\vartheta}}[\hat{\vartheta}'(W_{\hat{\vartheta}}(1) + \hat{\vartheta}) < \|\hat{\vartheta}\|^2/2 - \log\hat{z}] \\
&= \mu[\hat{\vartheta}' W(1) < -\|\hat{\vartheta}\|^2/2 - \log\hat{z}] = \Phi\left(-\frac{\log\hat{z} + \|\hat{\vartheta}\|^2/2}{\|\hat{\vartheta}\|}\right), \qquad (5.8)
\end{aligned}
$$

provided that the vector $\hat{\vartheta} \in K$ of (5.2) satisfies

$$\hat{\vartheta}'(\vartheta - \hat{\vartheta}) \geqslant 0, \qquad \forall \vartheta \in K. \qquad (5.9)$$

In this case, $\hat{z}$ is again given by

$$\log\frac{1}{\hat{z}} = \frac{1}{2}\|\hat{\vartheta}\|^2 + \|\hat{\vartheta}\| \cdot \Phi^{-1}(\alpha). \qquad (5.5)$$

Example 5.3 *Composite hypothesis* $\mathbf{Q} = \{\mu_\theta, \theta \in \Theta_1\}$ *versus composite alternative* $\mathbf{P} = \{\mu_\theta, \theta \in \Theta_2\}$.
Here Θ_i is the space of $\mathbf{F}$-progressively measurable processes $\theta\colon [0, 1] \times \Omega \to K_i$, $i = 1, 2$, where K_1 and K_2 are closed, convex subsets of $\mathbb{R}^d$ with $K_1 \cap K_2 = \varnothing$. We shall assume that there exist vectors $\hat{\vartheta}_1 \in K_1$, $\hat{\vartheta}_2 \in K_2$ such that

$$
\begin{aligned}
&\inf_{\substack{\vartheta_1\in K_1\\ \vartheta_2\in K_2}} \|\vartheta_2 - \vartheta_1\| = \|\hat{\vartheta}_2 - \hat{\vartheta}_1\|, \\
&(\hat{\vartheta}_2 - \hat{\vartheta}_1)'(\vartheta_1 - \hat{\vartheta}_1) \leqslant 0, \qquad \forall \vartheta_1 \in K_1.
\end{aligned} \qquad (5.10)
$$

In this case $H_{\theta_1} = Z_{\theta_1}(1)$ for $\theta_1(\cdot) \in \Theta_1$, and $G_{\theta_2} = Z_{\theta_2}(1)$ for $\theta_2(\cdot) \in \Theta_2$; both sets $\{H_\theta\}_{\theta\in\Theta_1}$, $\{G_\theta\}_{\theta\in\Theta_2}$ are convex and closed under μ-a.e. convergence, so the auxiliary dual problem of (3.11) has value function

$$\tilde{V}(z) = \inf_{\substack{\theta_1(\cdot)\in\Theta_1\\ \theta_2(\cdot)\in\Theta_2}} \mathrm{E}^{\mu}(Z_{\theta_2}(1) - zZ_{\theta_1}(1))^+ = \inf_{\substack{\theta_1(\cdot)\in\Theta_1\\ \theta_2(\cdot)\in\Theta_2}} \mathrm{E}^{\mu_{\theta_1}}\left(\frac{Z_{\theta_2}(1)}{Z_{\theta_1}(1)} - z\right)^+$$

$$= \inf_{\substack{\theta_1(\cdot)\in\Theta_1\\ \theta_2(\cdot)\in\Theta_2}} \mathrm{E}^{\mu_{\theta_1}}\left(\mathrm{e}^{\int_0^1(\theta_2(s)-\theta_1(s))'\,\mathrm{d}W_{\theta_1}(s)-\frac{1}{2}\int_0^1\|\theta_2(s)-\theta_1(s)\|^2\,\mathrm{d}s} - z\right)^+$$

$$= \mathrm{E}^{\mu_{\hat{\vartheta}_1}}\left(\mathrm{e}^{(\hat{\vartheta}_2-\hat{\vartheta}_1)'W_{\hat{\vartheta}_1}(1)-\frac{1}{2}\|\hat{\vartheta}_2-\hat{\vartheta}_1\|^2} - z\right)^+$$

$$= \int_{\mathbb{R}^d}\left(\mathrm{e}^{(\hat{\vartheta}_2-\hat{\vartheta}_1)'\xi-\frac{1}{2}\|\hat{\vartheta}_2-\hat{\vartheta}_1\|^2} - z\right)^+ \frac{\mathrm{e}^{-\frac{1}{2}\|\xi\|^2}}{(2\pi)^{d/2}}\,\mathrm{d}\xi$$

by analogy with (5.3), thanks to the first assumption of (5.10). Again, if we denote by $\hat{z}$ the unique $\arg\min_{0<z<\infty}[\alpha z + \tilde{V}(z)]$, the max-min optimal test of Theorem 4.1 has the form

$$\hat{X} = \mathbf{1}_{\{Z_{\hat{\vartheta}_2}(1)>\hat{z}Z_{\hat{\vartheta}_1}(1)\}} = 1_{\{(\hat{\vartheta}_2-\hat{\vartheta}_1)'W(1)>\frac{1}{2}(\|\hat{\vartheta}_2\|^2-\|\hat{\vartheta}_1\|^2)+\log\hat{z}\}}. \tag{5.11}$$

This test rejects the hypothesis if the inner product of $W(1)$ with $\hat{\vartheta}_2$ is sufficiently larger than its inner product with $\hat{\vartheta}_1$. Under the second assumption of (5.10), we also have

$$\alpha = \sup_{\theta_1(\cdot)\in\Theta_1} \mathrm{E}^{\mu_{\theta_1}}(\hat{X}) = \sup_{\theta_1(\cdot)\in\Theta_1} \mu_{\theta_1}[(\hat{\vartheta}_2-\hat{\vartheta}_1)'W(1) > (\|\hat{\vartheta}_2\|^2-\|\hat{\vartheta}_1\|^2)/2+\log\hat{z}]$$

$$= \sup_{\theta_1(\cdot)\in\Theta_1} \mu_{\theta_1}[(\hat{\vartheta}_2-\hat{\vartheta}_1)'W_{\theta_1}(1) + (\hat{\vartheta}_2-\hat{\vartheta}_1)'\int_0^1\theta_1(s)\,\mathrm{d}s > \tfrac{1}{2}(\|\hat{\vartheta}_2\|^2-\|\hat{\vartheta}_1\|^2)+\log\hat{z}]$$

$$= \mu_{\hat{\vartheta}_1}[(\hat{\vartheta}_2-\hat{\vartheta}_1)'W_{\hat{\vartheta}_1}(1) + (\hat{\vartheta}_2-\hat{\vartheta}_1)'\hat{\vartheta}_1 > (\|\hat{\vartheta}_2\|^2-\|\hat{\vartheta}_1\|^2)/2+\log\hat{z}]$$

$$= \mu_{\hat{\vartheta}_1}[(\hat{\vartheta}_2-\hat{\vartheta}_1)'W_{\hat{\vartheta}_1}(1) > \|\hat{\vartheta}_2-\hat{\vartheta}_1\|^2/2+\log\hat{z}], \tag{5.12}$$

from which we deduce

$$\log\frac{1}{\hat{z}} = \frac{1}{2}\|\hat{\vartheta}_1-\hat{\vartheta}_2\|^2 + \|\hat{\vartheta}_1-\hat{\vartheta}_2\|\cdot\Phi^{-1}(\alpha). \tag{5.13}$$

For our next few examples, let us switch to a new setting with Gaussian probability measures

$$\nu(\mathrm{d}\xi) = (2\pi)^{-d/2}\mathrm{e}^{-\frac{1}{2}\|\xi\|^2}\,\mathrm{d}\xi, \qquad \nu_\vartheta(\mathrm{d}\xi) = (2\pi)^{-\mathrm{d}/2}\mathrm{e}^{-\frac{1}{2}\|\xi-\vartheta\|^2}\,\mathrm{d}\xi \tag{5.14}$$

on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$; here again K is a compact, convex subset of $\mathbb{R}^d$ with $0\notin K$ and such that (5.2) holds. We shall denote by W the identity mapping on $\mathbb{R}^d$.

Example 5.4 *Composite hypothesis* $\mathbf{Q} = \{\nu_\vartheta, \vartheta\in K\}$ *versus simple alternative* $\mathbf{P} = \{\nu\}$. In this new setting, we have $G = 1$, $H_\vartheta = \mathrm{d}\nu_\vartheta/\mathrm{d}\nu = \mathrm{e}^{\vartheta'W-\frac{1}{2}\|\vartheta\|^2}$ for every $\vartheta\in K$. From Remark 4.1, the auxiliary dual problem of (3.11) now has value function

$$
\begin{aligned}
\tilde{V}(z) &= \inf_{\substack{\vartheta_1,\vartheta_2\in K\\ 0\leqslant\lambda\leqslant 1}} \mathrm{E}^{\nu}(1 - z[\lambda H_{\vartheta_1} + (1-\lambda)H_{\vartheta_2}])^+ \\
&= \inf_{\substack{\vartheta_1,\vartheta_2\in K\\ 0\leqslant\lambda\leqslant 1}} \mathrm{E}^{\nu}\Big(1 - z\Big[\lambda \mathrm{e}^{\vartheta_1' W - \frac{1}{2}\|\vartheta_1\|^2} + (1-\lambda)\mathrm{e}^{\vartheta_2' W - \frac{1}{2}\|\vartheta_2\|^2}\Big]\Big)^+ \\
&\geqslant \inf_{\theta(\cdot)\in\Theta} \mathrm{E}^{\mu}\Big(1 - z \cdot \mathrm{e}^{\int_0^1 \theta'(s)\,\mathrm{d}W(s) - \frac{1}{2}\int_0^1 \|\theta(s)\|^2\,\mathrm{d}s}\Big)^+ \\
&= \mathrm{E}^{\mu}\Big(1 - z\cdot \mathrm{e}^{\hat{\vartheta}' W(1) - \frac{1}{2}\|\hat{\vartheta}\|^2}\Big)^+ = \mathrm{E}^{\nu}\Big(1 - z\cdot \mathrm{e}^{\hat{\vartheta}' W - \frac{1}{2}\|\hat{\vartheta}\|^2}\Big)^+,
\end{aligned}
\tag{5.15}
$$

thanks to Example 5.2. Thus, the first infimum in (5.15) is attained by, say, $\vartheta_1 = \hat{\vartheta}$, $\vartheta_2 \in K$ arbitrary, $\lambda = 1$, and we have

$$
\tilde{V}(z) = \mathrm{E}^{\nu}\Big(1 - z\cdot \mathrm{e}^{\hat{\vartheta}' W - \frac{1}{2}\|\hat{\vartheta}\|^2}\Big)^+ = \int_{\mathbb{R}^d} \Big(1 - z\cdot \mathrm{e}^{\hat{\vartheta}'\xi - \frac{1}{2}\|\hat{\vartheta}\|^2}\Big)^+ \frac{\mathrm{e}^{-\frac{1}{2}\|\xi\|^2}}{(2\pi)^{d/2}}\,\mathrm{d}\xi
$$

as in (5.6). The test

$$
\hat{X} = \mathbf{1}_{\{\hat{z} H_{\hat{\vartheta}} < 1\}} = \mathbf{1}_{\{\hat{\vartheta}' W < \frac{1}{2}\|\hat{\vartheta}\|^2 - \log \hat{z}\}}
$$

of (5.7) rejects the hypothesis if the inner product of W with the vector $\hat{\vartheta} \in K$ of (5.2) is sufficiently small. Here $\hat{z}$ is once again given by (5.5).

Example 5.5 *Composite hypothesis* $\mathbf{Q} = \{\nu_\vartheta, \vartheta \in K_1\}$ *versus composite alternative* $\mathbf{P} = \{\nu_\vartheta, \vartheta \in K_2\}$.
Here K_1, K_2 are two disjoint, closed and convex subsets of $\mathbb{R}^d$ that satisfy the conditions of (5.10) for some $\hat{\vartheta}_1 \in K_1$, $\hat{\vartheta}_2 \in K_2$. (For instance, these conditions hold if $d = 1$ and if K_1, K_2 are two disjoint, closed intervals of the real line.) It is easy to see from the analysis of Example 5.3, that the test

$$
\hat{X} = \mathbf{1}_{\{(\hat{\vartheta}_2 - \hat{\vartheta}_1) W > \frac{1}{2}(\|\hat{\vartheta}_2\|^2 - \|\hat{\vartheta}_1\|^2) + \log \hat{z}\}} \tag{5.11}
$$

is max-min optimal, where

$$
\hat{z} = \exp\{-\|\hat{\vartheta}_1 - \hat{\vartheta}_2\|^2/2 - \Phi^{-1}(\alpha)\cdot\|\hat{\vartheta}_1 - \hat{\vartheta}_2\|\} \tag{5.13}
$$

attains the infimum of $z \mapsto \alpha z + \tilde{V}(z)$, and

$$
\tilde{V}(z) = \int_{\mathbb{R}^d} \Big(\mathrm{e}^{(\hat{\vartheta}_2 - \hat{\vartheta}_1)'\xi - \frac{1}{2}\|\hat{\vartheta}_2 - \hat{\vartheta}_1\|^2} - z\Big)^+ \frac{\mathrm{e}^{-\frac{1}{2}\|\xi\|^2}}{(2\pi)^{d/2}}\,\mathrm{d}\xi
$$

is the value function of the auxiliary dual problem (3.11).

Example 5.6 *Composite hypothesis* $\mathbf{Q} = \{\nu_\vartheta, \nu_{-\vartheta}\}$ *versus simple alternative* $\mathbf{P} = \{\nu\}$.
Here $\vartheta \in \mathbb{R}^d \backslash \{0\}$ is given. From Remark 4.1, the value function is now

$$\tilde{V}(z) = \inf_{0\leqslant\lambda\leqslant 1} \mathrm{E}^{\nu}\Big(1 - z\Big[\lambda \mathrm{e}^{\vartheta' W - \frac{1}{2}\|\vartheta\|^2} + (1-\lambda)\mathrm{e}^{-\vartheta' W - \frac{1}{2}\|\vartheta\|^2}\Big]\Big)^+$$

$$= \inf_{|\delta|\leqslant\frac{1}{2}} \mathrm{E}^{\nu}\left(1 - z\mathrm{e}^{-\frac{1}{2}\|\vartheta\|^2}\left[\left(\frac{1}{2}-\delta\right)\mathrm{e}^{\vartheta' W} + \left(\frac{1}{2}+\delta\right)\mathrm{e}^{-\vartheta' W}\right]\right)^+ = \inf_{|\delta|\leqslant\frac{1}{2}} f(z;\delta),$$

where

$$f(z;\delta) := \mathrm{E}^{\nu}\Big(1 + z\mathrm{e}^{-\frac{1}{2}\|\vartheta\|^2}[2\delta\sinh(\vartheta' W) - \cosh(\vartheta' W)]\Big)^+.$$

For every fixed $z \in (0,\infty)$, the function $\delta \mapsto f(z;\delta)$ is symmetric and attains its infimum over $[-\frac{1}{2},\frac{1}{2}]$ at $\delta = 0$, so that

$$\tilde{V}(z) = f(z;0) = \mathrm{E}^{\nu}\Big(1 - z\mathrm{e}^{-\frac{1}{2}\|\vartheta\|^2}\cosh(\vartheta' W)\Big)^+. \tag{5.16}$$

The optimal max-min test of (5.7), namely

$$\hat{X} = \mathbf{1}_{\left\{\hat{z}\left(\mathrm{e}^{\vartheta' W - \frac{1}{2}\|\vartheta\|^2} + \mathrm{e}^{-\vartheta' W - \frac{1}{2}\|\vartheta\|^2}\right) < 2\right\}} = \mathbf{1}_{\left\{\hat{z}\mathrm{e}^{-\frac{1}{2}\|\vartheta\|^2}\cosh(\vartheta' W) < 1\right\}} = \mathbf{1}_{\{|\vartheta' W| < h\|\vartheta\|\}}, \tag{5.17}$$

rejects the hypothesis if the absolute value of the inner product $\vartheta' W$ is sufficiently small. Here we have set

$$h := \frac{1}{\|\vartheta\|}\cosh^{-1}\left(\frac{\exp(\frac{1}{2}\|\vartheta\|^2)}{\hat{z}}\right), \tag{5.18}$$

where $\cosh^{-1}(\cdot)$ is the inverse of $\cosh(\cdot)$ on $(0,\infty)$ with $\cosh^{-1}(u) \equiv 0$ for $0 \leqslant u \leqslant 1$, and $\hat{z} = \arg\min_{0<z<\infty}[\alpha z + \tilde{V}(z)]$. The quantity $h > 0$ of (5.18) is characterized by

$$\alpha = \nu_{\pm\theta}(|\vartheta' W| < h\|\vartheta\|) = \nu(|\vartheta'(W \pm \vartheta)| < h\|\vartheta\|) = \nu(|\Xi \pm \|\vartheta\|\,| < h),$$

where $\Xi := \vartheta' W/\|\vartheta\|$ is standard normal under ν, or equivalently by the equation

$$\Phi(\|\vartheta\| + h) - \Phi(\|\vartheta\| - h) = \alpha. \tag{5.19}$$

Remark 5.1. It is not hard to verify that the test of (5.17)–(5.19) is also max-min optimal for testing the composite hypothesis $\mathbf{Q} = \{\nu_m,\ m \in (-\infty, -\vartheta] \cup [\vartheta, \infty)\}$ versus the simple alternative $\mathbf{P} = \{\nu\}$, for some $\vartheta > 0$, in the case $d = 1$.

6. Proofs

***Proof of* (3.2).** The *convexity* of the set (3.1) is obvious; now if $\{H_n\}_{n\in\mathbf{N}} \subseteq \mathscr{H}$ and $\lim_n H_n = H$, μ-a.e., then clearly $H \geqslant 0$, μ-a.e., and

$$\mathrm{E}^{\mu}(HX) = \mathrm{E}[\lim_n(H_n X)] \leqslant \liminf_n \mathrm{E}(H_n X) \leqslant \alpha$$

for every $X \in \mathscr{X}_\alpha$ by Fatou's lemma, so $H \in \mathscr{H}$. In other words, $\mathscr{H}$ is *closed under* μ-a.e. *convergence*. On the other hand, every H of the form $H_Q = \mathrm{d}Q/\mathrm{d}\mu$, for some $Q \in \mathbf{Q}$, belongs to $\mathscr{H}$, so that (3.2) holds as well, since $\mathrm{E}^\mu(H_Q X) = \mathrm{E}^Q(X) \leqslant \alpha$, $\forall X \in \mathscr{X}_\alpha$, from (2.4). Finally, from (2.4), the degenerate random variable $X \equiv \alpha$ belongs to $\mathscr{X}_\alpha$; with this choice in (3.1) we see that $\mathrm{E}^\mu(H) \leqslant 1$, $\forall H \in \mathscr{H}$, so that $\mathscr{H}$ is a *bounded* subset of $L^1(\mu)$. □

Proof of Proposition 3.1. From (3.4), (3.5) and (2.8) we have

$$\mathrm{E}^{\hat{P}}(X) \leqslant \alpha\hat{z} + \mathrm{E}^\mu(\hat{G} - \hat{z}\hat{H})^+, \qquad \forall X \in \mathscr{X}_\alpha,$$

$$\mathrm{E}^{\hat{P}}(\hat{X}) = \hat{z} \cdot \mathrm{E}^\mu(\hat{H}\hat{X}) + \mathrm{E}^\mu[(\hat{G} - \hat{z}\hat{H})\hat{X}] = \alpha\hat{z} + \mathrm{E}^\mu(\hat{G} - \hat{z}\hat{H})^+.$$

This leads to the first inequality in (3.8). On the other hand, (3.4) and (3.5) also give

$$\begin{aligned}\mathrm{E}^P(\hat{X}) &= \hat{z} \cdot \mathrm{E}^\mu(\hat{H}\hat{X}) + \mathrm{E}^\mu[\hat{X}(G_P - \hat{z}\hat{H})] \\ &= \alpha\hat{z} + \mathrm{E}^\mu[\hat{X}(G_P - \hat{z}\hat{H})], \qquad \forall P \in \mathbf{P}.\end{aligned}$$

Thanks to assumption (3.7), this last quantity dominates

$$\alpha\hat{z} + \mathrm{E}^\mu[\hat{X}(\hat{G} - \hat{z}\hat{H})] = \mathrm{E}^{\hat{P}}(\hat{X}),$$

and the second inequality of (3.8) follows. □

Proof of Proposition 3.2. From (3.4) we have

$$\mathrm{E}^\mu(G - zH)^+ + \alpha z \geqslant \mathrm{E}^P(X), \qquad \forall\, (G, H) \in \mathscr{G} \times \mathscr{H}, \tag{6.1}$$

for every $z > 0$, $X \in \mathscr{X}_\alpha$ and with $P = \int G\, \mathrm{d}\mu$. On the other hand, (3.6) gives

$$\mathrm{E}^\mu(\hat{G} - \hat{z}\hat{H})^+ + \alpha\hat{z} = \mathrm{E}^{\hat{P}}(\hat{X}). \tag{3.6$'$}$$

Now substitute $z = \hat{z}$, $X = \hat{X}$ into (6.1) to obtain (i), in conjunction with (3.6)$'$ and the second inequality of (3.8). Similarly, substitute $z > 0$ arbitrary, and $X \in \mathscr{X}_\alpha$, into (6.1) to obtain (ii) and $V_*(\alpha) = \mathrm{E}^{\hat{P}}(\hat{X})$, again in conjunction with (3.6)$'$ and the second inequality of (3.8). Property (iv) then follows from Proposition 3.1, and (iii) is an easy consequence of (i), (ii). □

Proof of Lemma 4.1. From $(G - z_1 H)^+ - (G - z_2 H)^+ \leqslant |z_1 - z_2| H$, from $\mathrm{E}^\mu(H) \leqslant 1$, and from (3.11), we obtain

$$\begin{aligned}\tilde{V}(z_1) &\leqslant \mathrm{E}^\mu(G - z_1 H)^+ \leqslant |z_1 - z_2| \cdot \mathrm{E}^\mu(H) + \mathrm{E}^\mu(G - z_2 H)^+ \\ &\leqslant \mathrm{E}^\mu(G - z_2 H)^+ + |z_1 - z_2|, \qquad \forall\, (G, H) \in \mathscr{G} \times \mathscr{H}.\end{aligned}$$

Taking the infimum over this set, we obtain $\tilde{V}(z_1) \leqslant \tilde{V}(z_2) + |z_1 - z_2|$, and then we simply interchange the roles of z_1 and z_2. □

Proof of Lemma 4.2. Let $\{(G_n, H_n)\}_{n \in \mathbf{N}} \subset \mathscr{G} \times \mathscr{H}$ be a minimizing sequence for (3.11). Because $\mathscr{G} \times \mathscr{H}$ is bounded in $L^1(\mu) \times L^1(\mu)$, there exists by the theorem of Komlós (1967) (see also Schwartz 1986) a random vector $(\hat{G}, \hat{H}) \in L^1(\mu) \times L^1(\mu)$ and a relabelled subsequence $\{(G'_j, H'_j)\}_{j\in\mathbf{N}} \subseteq \{(G_n, H_n)\}_{n\in\mathbf{N}}$, such that

$$(\tilde{G}_k, \tilde{H}_k) := \left(\frac{1}{k}\sum_{j=1}^{k} G'_j, \frac{1}{k}\sum_{j=1}^{k} H'_j\right) \to (\hat{G}, \hat{H})$$

as $k \to \infty$, μ-a.e. By the convexity and the μ-a.e. closedness of $\mathscr{G}$ and $\mathscr{H}$, we have $(\hat{G}, \hat{H}) \in \mathscr{G} \times \mathscr{H}$; on the other hand, Fatou's lemma and convexity give

$$\begin{aligned}\mathrm{E}^{\mu}(\hat{G} - z\hat{H})^{+} &= \mathrm{E}^{\mu}[\lim_k(\tilde{G}_k - z\tilde{H}_k)^{+}] \\ &\leqslant \liminf_k \mathrm{E}^{\mu}(\tilde{G}_k - z\tilde{H}_k)^{+} = \liminf_k \mathrm{E}^{\mu}\left(\frac{1}{k}\sum_{j=1}^{k}(G'_j - zH'_j)\right)^{+} \\ &\leqslant \liminf_k \mathrm{E}^{\mu}\left(\frac{1}{k}\sum_{j=1}^{k}(G'_j - zH'_j)^{+}\right) = \lim_j \mathrm{E}^{\mu}(G'_j - zH'_j)^{+} = \tilde{V}(z).\end{aligned}$$

□

Proof of Lemma 4.3. The convex function

$$f_\alpha(z) := \alpha z + \tilde{V}(z) = \alpha z + \inf_{(G,H)\in(\mathscr{G}\times\mathscr{H})} \mathrm{E}^{\mu}(G - zH)^{+}, \qquad 0 < z < \infty,$$

satisfies $f_\alpha(0+) = \inf_{G\in\mathscr{G}} \mathrm{E}^{\mu}(G) = 1$, $f_\alpha(z) \geqslant \alpha z$ and thus $f_\alpha(\infty) = \infty$. Consequently, $f_\alpha(\cdot)$ either attains its infimum at some $\hat{z}_\alpha \in (0, \infty)$ as claimed, or else satisfies $f_\alpha(z) \geqslant 1$, $\forall z > 0$. This latter possibility can easily be ruled out, as it implies

$$\alpha z \geqslant 1 - \tilde{V}(z) \geqslant 1 - \mathrm{E}^{\mu}(G - zH)^{+} = \mathrm{E}^{\mu}(G) - \mathrm{E}^{\mu}(G - zH)^{+} \geqslant z \cdot \mathrm{E}^{\mu}[H 1_{\{zH \leqslant G\}}], \qquad \forall z > 0,$$

for any given $G \in \mathscr{G}$, $H \in \mathscr{H}$; dividing by z and then letting $z \downarrow 0$, we obtain $\mathrm{E}^{\mu}(H) \leqslant \alpha$ for every $H \in \mathscr{H}$, and thus $\alpha \geqslant 1$, a contradiction. □

Proof of Lemma 4.4. The convexity of $\mathscr{G}$ and $\mathscr{H}$ leads to that of the set $\mathscr{M}$ in (4.2). Now consider a sequence $\{(G_n, H_n, z_n)\}_{n\in\mathbf{N}} \subseteq \mathscr{M}$ which converges to some $(G, H, z) \in \mathscr{L}$ in the norm $\|\cdot\|$ of Lemma 4.4. Then $z_n \to z$ and $(G_{n_j}, H_{n_j}) \to (G, H)$, μ-a.e. (along a subsequence), so that $G \in \mathscr{G}$ and $H \in \mathscr{H}$ by the closedness of both $\mathscr{G}$ and $\mathscr{H}$ under μ-a.e. convergence. The properness, convexity and lower-semicontinuity of the functional (4.3) on $\mathscr{L}$ are relatively easy to check. For the final claim of the lemma, just observe from (3.13), (3.11) that we have

$$\alpha\hat{z} + \mathrm{E}^{\mu}(\hat{G} - \hat{z}\hat{H})^{+} = \alpha\hat{z} + \tilde{V}(\hat{z}) \leqslant \alpha z + \tilde{V}(z) \leqslant \alpha z + \mathrm{E}^{\mu}(G - zH)^{+}, \qquad \forall\, (G, H, z) \in \mathscr{M}.$$

□

Proof of Theorem 4.1. First, let us read (4.7) with

$$G := \hat{G} - \mathbf{1}_{\{\hat{X}+\hat{W}>0\}}, \qquad Y := \hat{z}\hat{H} - \mathbf{1}_{\{\hat{X}+\hat{W}>0\}},$$

and in conjunction with (4.8), to obtain $\mathrm{E}^{\mu}(\hat{X}+\hat{W})^{+}=\mathrm{E}^{\mu}[(\hat{X}+\hat{W})\mathbf{1}_{\{\hat{X}+\hat{W}\geqslant 0\}}]\leqslant 0$. Similarly, we can derive $\mathrm{E}^{\mu}(\hat{X}+\hat{W})^{-}\leqslant 0$, whence $\hat{X}+\hat{W}=0$, μ-a.e. This proves the first equality in (4.9) and shows, in conjunction with (4.8), that the conditions of (4.6), (4.7) can be written as

$$\mathrm{E}^{\mu}[\hat{X}(\hat{z}\hat{H}-zH+G-\hat{G})]\geqslant\alpha(\hat{z}-z),\qquad\forall\,(G,H,z)\in\mathscr{M},\tag{4.6$'$}$$

$$\mathrm{E}^{\mu}[\hat{X}(Y-\hat{z}\hat{H}+\hat{G}-G)+(G-Y)^{+}-(\hat{G}-\hat{z}\hat{H})^{+}]\geqslant 0,\qquad\forall\,(G,Y)\in(L^{1}(\mu))^{2}.\tag{4.7$'$}$$

From now on we shall write

$$\hat{X}=\mathbf{1}_{\{\hat{z}\hat{H}<\hat{G}\}}+A\tag{6.2}$$

for a suitable random variable A. Substituting $G\equiv\hat{G}$, and $\hat{X}$ as in (6.2), into (4.7)$'$, we obtain

$$\begin{aligned}\mathrm{E}^{\mu}[(\hat{z}\hat{H}-Y)A]&\leqslant\mathrm{E}^{\mu}[(\hat{G}-Y)^{+}-(\hat{G}-\hat{z}\hat{H})\mathbf{1}_{\{\hat{z}\hat{H}<\hat{G}\}}+(Y-\hat{z}\hat{H})\mathbf{1}_{\{\hat{z}\hat{H}<\hat{G}\}}]\\&=\mathrm{E}^{\mu}[(\hat{G}-Y)\mathbf{1}_{\{Y<\hat{G}\}}+(Y-\hat{G})\mathbf{1}_{\{\hat{z}\hat{H}<\hat{G}\}}],\qquad\forall\,Y\in L^{1}(\mu).\end{aligned}\tag{4.7$''$}$$

Now take $Y\in L^{1}(\mu)$ such that

$$\{Y<\hat{G}\}=\{\hat{z}\hat{H}<\hat{G}\}\qquad\text{mod }\mu,\tag{6.3}$$

which implies

$$\mathrm{E}^{\mu}[A(\hat{z}\hat{H}-Y)]\leqslant 0,\qquad\text{for every }Y\in L^{1}(\mu)\text{ as in (6.3)},\tag{4.7$'''$}$$

as well as

$$A\leqslant 0,\qquad\mu\text{-a.e. on }\{\hat{z}\hat{H}<\hat{G}\},\tag{6.4}$$

$$A\geqslant 0,\qquad\mu\text{-a.e. on }\{\hat{z}\hat{H}\geqslant\hat{G}\}.\tag{6.5}$$

To prove (6.4), let us assume $\mu(A>0,\,\hat{z}\hat{H}<\hat{G})>0$; we can select $Y\in L^{1}(\mu)$ negative and with $|Y|$ so large on this set as to make $\mathrm{E}^{\mu}[A(\hat{z}\hat{H}-Y)\mathbf{1}_{\{A>0,\hat{z}\hat{H}<\hat{G}\}}]$ positive and large enough to violate (4.7)$'''$ written in the form

$$\mathrm{E}^{\mu}[A(\hat{z}\hat{H}-Y)\mathbf{1}_{\{A>0,\hat{z}\hat{H}<\hat{G}\}}]+\mathrm{E}^{\mu}[A(\hat{z}\hat{H}-Y)\mathbf{1}_{\{A\leqslant 0,\hat{z}\hat{H}<\hat{G}\}}]+\mathrm{E}^{\mu}[A(\hat{z}\hat{H}-Y)\mathbf{1}_{\{\hat{z}\hat{H}\geqslant\hat{G}\}}]\leqslant 0.$$

Similarly, in order to prove (6.5), assume that the set $\{A<0,\,\hat{z}\hat{H}\geqslant\hat{G})$ has positive μ-probability, and select $Y\in L^{1}(\mu)$ positive and so large on this set as to violate condition (4.7)$'''$.

We can also see that

$$A=0,\qquad\mu\text{-a.e. on }\{\hat{z}\hat{H}<\hat{G}\},\tag{6.6}$$

$$A=0,\qquad\mu\text{-a.e. on }\{\hat{z}\hat{H}>\hat{G}\}.\tag{6.7}$$

To prove (6.6), assume $\mu(A<0,\,\hat{z}\hat{H}<\hat{G})>0$ so that $\delta:=\mathrm{E}^{\mu}[A(\hat{z}\hat{H}-\hat{G})\mathbf{1}_{\{\hat{z}\hat{H}<\hat{G}\}}]>0$ in conjunction with (6.4). For arbitrary $\varepsilon>0$, the random variable

$$Y := \begin{cases} \hat{G} - \varepsilon, & \text{on } \{\hat{z}\hat{H} < \hat{G}\}, \\ \hat{G}, & \text{on } \{\hat{z}\hat{H} \geqslant \hat{G}\}, \end{cases} \tag{6.8}$$

is in $L^1(\mu)$, satisfies (6.3), and for it the condition (4.7)‴ becomes

$$\begin{aligned} 0 &\geqslant \mathrm{E}^{\mu}\Big[A(\hat{z}\hat{H} - \hat{G})\mathbf{1}_{\{\hat{z}\hat{H} \geqslant \hat{G}\}}\Big] + \mathrm{E}^{\mu}\Big[A(\hat{z}\hat{H} - \hat{G} + \varepsilon)\mathbf{1}_{\{\hat{z}\hat{H} < \hat{G}\}}\Big] \\ &= \mathrm{E}^{\mu}\Big[A(\hat{z}\hat{H} - \hat{G})\mathbf{1}_{\{\hat{z}\hat{H} \geqslant \hat{G}\}}\Big] + \delta + \varepsilon \cdot \mathrm{E}^{\mu}\Big[A\mathbf{1}_{\{\hat{z}\hat{H} < \hat{G}\}}\Big]. \end{aligned}$$

Letting $\varepsilon \downarrow 0$, this gives $\delta + \mathrm{E}^{\mu}[A(\hat{z}\hat{H} - \hat{G})^+] \leqslant 0$, which is absurd in light of (6.5). This proves the property (6.6), and allows us to recast (4.7)‴ in the equivalent form

$$\mathrm{E}^{\mu}\Big[A(\hat{z}\hat{H} - Y)\mathbf{1}_{\{\hat{z}\hat{H} \geqslant \hat{G}\}}\Big] \leqslant 0, \qquad \text{for every } Y \in L^1(\mu) \qquad \text{as in (6.3).}$$

In particular, with Y as in (6.8), this gives $\mathrm{E}^{\mu}[A(\hat{z}\hat{H} - \hat{G})^+] \leqslant 0$ which, in light of (6.5), leads to (6.7).

The properties (6.4)–(6.7) show that we have

$$A = B \cdot \mathbf{1}_{\{\hat{z}\hat{H} = \hat{G}\}} \qquad \text{for some random variable } B \geqslant 0. \tag{6.9}$$

Using this in (4.7)″ gives

$$\mathrm{E}^{\mu}[(\hat{G} - Y)B\mathbf{1}_{\{\hat{z}\hat{H} = \hat{G}\}}] \leqslant \mathrm{E}^{\mu}[(\hat{G} - Y)\mathbf{1}_{\{Y < \hat{G}\}} + (Y - \hat{G})\mathbf{1}_{\{\hat{z}\hat{H} < \hat{G}\}}], \qquad \forall\, Y \in L^1(\mu). \tag{6.10}$$

Finally, we have to establish

$$0 \leqslant B \leqslant 1 \qquad \mu\text{-a.e. on } \{\hat{z}\hat{H} = \hat{G}\}.$$

In order to show this, we shall substitute

$$Y := \begin{cases} 0, & \text{on } \{B > 1, \hat{z}\hat{H} = \hat{G}\}, \\ \hat{G}, & \text{on } \{B \leqslant 1, \hat{z}\hat{H} = \hat{G}\}, \\ \hat{G}(1 - \varepsilon), & \text{otherwise,} \end{cases}$$

for arbitrary $0 < \varepsilon < 1$, into (6.10); this leads to

$$\mathrm{E}^{\hat{P}}\Big[B \cdot \mathbf{1}_{\{B > 1, \hat{z}\hat{H} = \hat{G}\}}\Big] \leqslant \hat{P}[B > 1, \hat{z}\hat{H} = \hat{G}] + \varepsilon \cdot \hat{P}[\hat{z}\hat{H} > \hat{G}]$$

for every $0 < \varepsilon < 1$, as well as to

$$\mathrm{E}^{\hat{P}}[(B - 1)^+ \cdot \mathbf{1}_{\{\hat{z}\hat{H} = \hat{G}\}}] \leqslant 0$$

in the limit as $\varepsilon \downarrow 0$. Therefore, $B \leqslant 1$ μ-a.e. on $\{\hat{z}\hat{H} = \hat{G}\}$. This concludes the proof that $\hat{X}$ is of the form (2.8), (4.9). In particular, $\mu(0 \leqslant \hat{X} \leqslant 1) = 1$ and (3.5b) holds, so $\hat{X} \in \mathscr{X}_a$. □

***Proof of* (5.3) and (5.6).** Here we follow Xu and Shreve (1992). Let $\varphi : \mathbb{R} \to [0, \infty)$ be a convex function satisfying a linear growth condition; with $\hat{\vartheta} \in K$ as in (5.2), and arbitrary $\theta(\cdot) \in \Theta$, we shall show that

$$\mathrm{E}^{\mu}\varphi(Z_{\theta}(t)) \geqslant \mathrm{E}^{\mu}\varphi(Z_{\hat{\vartheta}}(t)), \qquad 0 \leqslant t \leqslant 1. \tag{6.11}$$

This will clearly prove (5.3) by taking $t = 1$ and $\varphi(x) = (x - z)^+$, and (5.6) by taking $t = 1$ and $\varphi(x) = (1 - zx)^+$. We can write

$$\frac{\theta(\cdot)}{\|\theta(\cdot)\|} = Q(\cdot)\frac{\hat{\vartheta}}{\|\hat{\vartheta}\|} \qquad \text{a.e. on } [0,\, 1] \times \Omega,$$

for some **F**-progressively measurable process $Q(\cdot)$ with values in the space of $d \times d$ orthonormal matrices. Let us also observe that the process

$$\Lambda(t) := \int_0^t \frac{\|\theta(s)\|^2}{\|\hat{\vartheta}\|^2}\, \mathrm{d}s \geqslant t, \qquad 0 \leqslant t \leqslant 1,$$

is increasing with $\mathrm{d}\Lambda(\cdot)/\mathrm{d}t \geqslant 1$, so that its inverse is well defined and $\Lambda^{-1}(u)$ is an **F**-stopping time with $\Lambda^{-1}(u) \leqslant u$, for every $0 \leqslant u \leqslant 1$. Now the time-changed process

$$\hat{W}(u) := \int_0^{\Lambda^{-1}(u)} \frac{\|\theta(s)\|}{\|\hat{\vartheta}\|} Q'(s)\, \mathrm{d}W(s), \qquad 0 \leqslant u \leqslant 1,$$

is a (local) martingale of the filtration $\hat{\mathscr{F}}(u) := \mathscr{F}(\Lambda^{-1}(u))$, $0 \leqslant u \leqslant 1$, with

$$\langle \hat{W}_i,\, \hat{W}_j \rangle(u) = \delta_{ij} \int_0^{\Lambda^{-1}(u)} \frac{\|\theta(s)\|^2}{\|\hat{\vartheta}\|^2}\, \mathrm{d}s = u \cdot \delta_{ij}.$$

In other words, $\hat{W}(\cdot)$ is an $\hat{\mathbf{F}}$-Brownian motion (by P. Lévy's Theorem 3.3.16 in Karatzas and Shreve 1991), and

$$\begin{aligned} 1 + \int_0^u Z_\theta(\Lambda^{-1}(\tau))\hat{\vartheta}'\, \mathrm{d}\hat{W}(\tau) &= 1 + \int_0^{\Lambda^{-1}(u)} Z_\theta(s)\hat{\vartheta}'\, \mathrm{d}\hat{W}(\Lambda(s)) \\ = 1 + \int_0^{\Lambda^{-1}(u)} Z_\theta(s) \frac{\|\theta(s)\|}{\|\hat{\vartheta}\|} (Q(s)\hat{\vartheta})'\, \mathrm{d}W(s) &= 1 + \int_0^{\Lambda^{-1}(u)} Z_\theta(s)\theta'(s)\, \mathrm{d}W(s) \\ = Z_\theta(\Lambda^{-1}(u)), \qquad 0 \leqslant u \leqslant 1. \end{aligned}$$

Since $Z_{\hat{\vartheta}}(t) = 1 + \int_0^t Z_{\hat{\vartheta}}(s)\hat{\vartheta}'\, \mathrm{d}W(s)$, $0 \leqslant t \leqslant 1$, we conclude that the processes $Z_{\hat{\vartheta}}(\cdot)$, $Z_\theta(\Lambda^{-1}(\cdot))$ have the same distribution. From this, from the optional sampling theorem, and from the fact that $\varphi(Z_\theta(\cdot))$ is a submartingale, we obtain

$$\mathrm{E}^\mu \varphi(Z_{\hat{\vartheta}}(t)) = \mathrm{E}^\mu \varphi(Z_\theta(\Lambda^{-1}(t))) \leqslant \mathrm{E}^\mu \varphi(Z_\theta(t)),$$

which proves (6.11). □

Acknowledgement

The research for this paper was supported in part by National Science Foundation grant NSF-DMS-97-32810. We are grateful to Professor Hans Föllmer for calling our attention to the book of Witting (1985), and to the referees and the editor for pointing out relevant literature.

References

Aubin, J. and Ekeland, I. (1984) *Applied Nonlinear Analysis*. New York: Wiley.

Baumann, V. (1968) Eine parameterfreie Theorie der ungünstigsten Verteilungen für das Testen von Hypothesen. *Z. Wahrscheinlichkeitstheorie Verw. Geb.*, **11**, 61–73.

Beneš, V.E. (1971) Existence of optimal stochastic control laws. *SIAM J. Control*, **8**, 446–475.

Cvitanić, J. (2000) Minimizing expected loss of hedging in incomplete and constrained markets. *SIAM J. Control Optim.* To appear.

Cvitanić, J. and Karatzas, I. (1999) On dynamic measures of risk. *Finance Stochastics*, **3**, 451–482.

Ferguson, T.S. (1967) *Mathematical Statistics: A Decision-Theoretic Approach*. New York and London: Academic Press.

Föllmer, H. and Leukert, P. (1999) Quantile hedging. *Finance Stochastics*, **3**, 251–273.

Föllmer, H. and Leukert, P. (2000) Partial hedging with minimal shortfall risk. *Finance Stochastics*, **4**, 117–146.

Heath, D. (1993) A continuous-time version of Kulldorff's result. Unpublished Manuscript, Cornell University.

Huber, P. and Strassen, V. (1973) Minimax tests and the Neyman–Pearson lemma for capacities. *Ann. Statist.*, **1**, 251–263.

Karatzas, I. (1997) Adaptive control of a diffusion to a goal, and a parabolic Monge–Ampère–type equation. *Asian J. Math.*, **1**, 295–313.

Karatzas, I. and Shreve, S.E. (1991) *Brownian Motion and Stochastic Calculus* (2nd edition). New York: Springer-Verlag.

Komlós, J. (1967) A generalization of a problem of Steinhaus. *Acta Math. Acad. Sci. Hungar.*, **18**, 217–229.

Krafft, O. and Witting, H. (1967) Optimale Tests und ungünstigste Verteilungen. *Z. Wahrscheinlichkeitstheorie Verw. Geb.*, **7**, 289–302.

Lehmann, E.L. (1952) On the existence of least-favorable distributions. *Ann. Math. Statist.*, **23**, 408–416.

Lehmann, E.L. (1986) *Testing Statistical Hypotheses* (2nd edition). New York: Wiley.

Schwartz, M. (1986) New proofs of a theorem of Komlós. *Acta Math. Hungar.*, **47**, 181–185.

Spivak, G. (1998) Maximizing the probability of perfect hedge. Doctoral dissertation, Department of Statistics, Columbia University.

Vajda, I. (1989) *Theory of Statistical Inference and Information.* Dordrecht: Kluwer.

Witting, H. (1985) *Mathematische Statistik I*. Stuttgart: B.G. Teubner.

Xu, G.-L. and Shreve, S.E. (1992) A duality method for optimal consumption and investment under short-selling prohibition. II: Constant market coefficients. *Ann. Appl. Probab.*, **2**, 314–328.

Received September 1999 and revised June 2000